



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

PCT/EP 03/07028

23 DEC 2003

54329  
10/519210

REC'D 16 SEP 2003

WIPO PCT

Bescheinigung

Certificate

Attestation

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

03008444.6

**PRIORITY DOCUMENT**  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OF (b)

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

R C van Dijk



Anmeldung Nr:  
Application no.: 03008444.6  
Demande no:

Anmeldetag:  
Date of filing: 11.04.03  
Date de dépôt:

Anmelder/Applicant(s)/Demandeur(s):

BASF AKTIENGESELLSCHAFT

67056 Ludwigshafen  
ALLEMAGNE

Bezeichnung der Erfindung/Title of the invention/Titre de l'invention:  
(Falls die Bezeichnung der Erfindung nicht angegeben ist, siehe Beschreibung.  
If no title is shown please refer to the description.  
Si aucun titre n'est indiqué se referer à la description.)

Plasmid vectors for transformation of filamentous fungi

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s)  
revendiquée(s)

Staat/Tag/Aktenzeichen/State/Date/File no./Pays/Date/Numéro de dépôt:

Internationale Patentklassifikation/International Patent Classification/  
Classification internationale des brevets:

C12N15/00

Am Anmeldetag benannte Vertragstaaten/Contracting states designated at date of  
filing/Etats contractants désignées lors du dépôt:

AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HU IE IT LU MC NL  
PT SE SI SK TR LI

## Plasmid vectors for transformation of filamentous fungi

## Description

5

The invention relates to novel plasmid vectors for transformation of filamentous fungi and to a method of modifying the genome of filamentous fungi based on these vectors. The invention furthermore relates to the modification of a specific gene via the process of homologous recombination, to recombinant expression of foreign genes in filamentous fungi and to new selection markers for detecting the successful transfer of the target gene in filamentous fungi.

15 One method currently used for transformation of filamentous fungi is random mutagenesis based on transposons insertion, a method also known for plant transformation (WO 01/38509). This method allows to study the genomes for several species such as Magnaporthe grisea (for examples WO 00/55346; WO 00/56902). However, this strategy requires a big effort in term of bioinformatic tools and molecular biology to localise precisely the insertion in the genome.

Alternatively known transformation methods are based on targeted integration. Targeted transformation of fungi can be carried out either by offering a Knock out cassette with a marker-gene flanked by two homologous sequences (Aronson et al, 1994, Mol. Gen. Genet. 242: 490-494; Royer et al, 1999, Fungal Genetics and Biology 28: 68-78; Schaefer, 2001, Current Opinion in Plant Biology 4: 143-150) or by quoting a plasmid with the marker gene in the neighbourhood of a homologous sequence (Shortle et al., 1982, Science 217: 371-373; Bird and Bradshaw, 1997, Mol Gen Genet. 255: 219-225; Feng et al., 2001, Infection and Immunity 69 (3): 1781-1794; Schaefer, 2001, Current Opinion in Plant Biology 4: 143-150). Both procedures are in principle attractive methods to study the gene function, but they have the disadvantage of a high frequency of integration at ectopic sites by illegitimate recombination. The gene targeting efficiency (gene targeting / gene targeting + illegitimate recombination) is 95% for *S. cerevisiae*, 10-90% for *S. pombe*, 5-75% for *Aspergillus nidulans* and 1-30% for *Neurospora crassa* using a size of homology of 2-9 Kb (Schaefer, 2001, Current Opinion in Plant Biology 4: 143-150). Especially for filamentous fungi this side effect is quite high, if conventional plasmid vectors are used.

45

In addition, the efficiency of the gene targeting increases if the length of homologous recombination region is rised (Schaefer, 2001, Current Opinion in Plant Biology 4: 143-150). Thus, plasmid vectors currently used comprise gene fragments of the gene to be  
5 knocked out of a size of at least 2000bp as indicated above. The overall size of these plasmid vectors is at least 8000bp (P. J. Punt et. al., 1992, Methods in Enzymology, vol.216, pp 447-457; ; Schaefer, 2001, Current Opinion in Plant Biology 4: 143-150). Since transformation efficiency is decreasing by increase of the  
10 plasmid vector size, transformation efficiency is unsatisfactory giving rise to long times until positive clones can be identified. This is an obstacle especially for large scale genomic analysis projects or recombinant expression.

15 Furthermore, currently used plasmid vectors contain many unique restriction sites, rising difficulties in construction of the knock-out (KO-) plasmids and the transformation process. The efficiency of homologous recombination is improved when the KO-plasmid is digested with a restriction enzyme presenting a unique  
20 site in the middle of DNA fragment homologous to the targeted gene. The presence of high amount of restriction sites especially unique ones in the plasmid backbone decrease the chance to find a natural restriction site in the appropriate location of the targeted DNA fragment. This problem is usually overcome by modification  
25 tion of the targeted DNA fragment requiring several cloning step and additional manipulation in terms of molecular biology, what is a disadvantageous time consuming methology.

Integration of recombinant gene by homologous recombination in  
30 fungi is also a tool to identify gene function for essential gene: the biochemical characterization of an essential gene cannot be studied by classical knock-out strategy since the mutants carrying a disruption of such a gene are not viable. One way consisting to overexpress such a gene overcome the problem when a  
35 typical phenotype can be assigned to the mutant that overexpresses the gene. Another approach can be to regulate the gene expression by an inducible promoter sequence so that the gene could be expressed or repressed when needed and consequently permits to isolate viable mutants. As mentioned above, these approaches re-  
40 quire at least several thousand bp of the nucleic acid sequence to be studied that need to be integrated in the genome of the fungi together with a plasmid vector comprising the different parts of the nucleic acid sequence. In addition, if the recombinant DNA is integrated at an ectopic site, the identification of  
45 the mutant strains becomes more complicated and the position of the integration in the genome may influence the level of expression of the recombinant protein. Taking the aforesaid into con-

## 3

sideration, currently existing plasmid vectors currently used for transformation of filamentous fungi exhibit a lot of disadvantages and are neither suitable for large scale analysis e.g. in functional genomic studies nor convenient for recombinant expression in a filamentous fungi. Additionally, there is a constant need for new selection markers facilitating the selection process.

Thus, object of the present invention was to develop tools for targeted transformation of filamentous fungi that overcome the disadvantages of the state of the art like plasmid vectors suitable for functional genomic studies and recombinant expression and new selection markers.

Surprisingly, we have found that the object of the invention has been achieved by construction of a plasmid vector for targeted transformation of filamentous fungi comprising

- a) an origin of replication for a host organism which is not originating from the filamentous fungi to be transformed;
  - b) a selection marker for a host organism not originating from the filamentous fungi to be transformed;
  - c) a promotor facilitating recombinant expression in filamentous fungi that is functionally linked to the coding region of the hygromycin resistance gene which is functionally linked to a terminator which facilitates transcription termination in filamentous fungi;
- wherein the overall size of the elements a), b) and c) does not exceed 4500 bp; and
- d) a nucleic acid sequence, which is homologous to nucleic acid sequences of the filamentous fungi to be transformed and makes homologous recombination in the filamentous fungi to be transformed possible.

The term overall size of the elements a), b) and c) designates the combination of the essential elements of the expression vector without the nucleic acid sequence d).

The overall size of the elements a), b) and c) does not exceed 4500 bp, preferably 4100 bp, more preferably 3700 bp.

In addition to the nucleic acid elements a), b), c) and d) the plasmid vector optionally comprises a cloning site containing rare restriction sites or a TA-cloning site by which further nucleic acid sequences can be cloned easily into the plasmid vector. A TA-cloning site comprises thymidine residues linked onto the 3'-ends of linearized plasmid DNA, which would allow some annealing to occur between the vector and the A-tailed PCR product to be ligated. This process is called TA cloning. Preferably, the vector is modified that there are only few unique restriction sites left enabling the digestion by commercial available restriction enzymes of the homologous sequence of the targeted gene prior to the transformation.

Filamentous fungi that can be transformed with the vectors of the present invention are non-phytopathogenic filamentous fungi e.g. *Neurospora* species like *Neurospora crassa* and phytopathogenic filamentous fungi, wherein the phytopathogenic filamentous fungi are preferred. Examples of other non-phytopathogenic filamentous fungi are *Aspergillus* species such as *Aspergillus parasiticus*, *Aspergillus nidulans*, *Aspergillus niger* and *Wangiella* such as *Wangiella dermatidis*. Preferred phytopathogenic filamentous fungi are selected from the group consisting of the genera *Alternaria*, *Podosphaera*, *Sclerotinia*, *Physalospora*, *Botrytis*, *Corynespora*, *Colletotrichum*, *Diplocarpon*, *Elsinoe*, *Diaporthe*, *Sphaerotheca*, *Cinula*, *Cercospora*, *Erysiphe*, *Sphaerotheca*, *Leveillula*, *Mycosphaerella*, *Phyllactinia*, *Gloesporium*, *Gymnosporangium*, *Leptotthrydium*, *Podosphaera*, *Gloedes*, *Cladosporium*, *Phomopsis*, *Phytopora*, *Phytophthora*, *Erysiphe*, *Fusarium*, *Verticillium*, *Glomerella*, *Drechslera*, *Bipolaris*, *Personospora*, *Phaeoisariopsis*, *Spaceloma*, *Pseudocercospora*, *Pseudoperonospora*, *Puccinia*, *Typhula*, *Pyricularia*, *Rhizoctonia*, *Stachosporium*, *Uncinula*, *Ustilago*, *Gaeumannomyces* and *Fusarium*, more preferred from the group consisting of the genera and species *Alternaria*, *Podosphaera*, *Sclerotinia*, *Physalospora* such as *Physalospora* canker, *Botrytis* species such as *Botrytis cinerea*, *Corynespora* such as *Corynespora melonis*, *Colletotrichum*, *Diplocarpon* such as *Diplocarpon rosae*, *Elsinoe* such as *Elsinoe fawcetti*, *Diaporthe* such as *Diaporthe citri*, *Sphaerotheca*, *Cinula* such as *Cinula neccata*, *Cercospora*, *Erysiphe* such as *Erysiphe cichoracearum* and *Erysiphe graminis*, *Sphaerotheca* such as *Sphaerotheca fuliginea*, *Leveillula* such as *Leveillula taurica*, *Mycosphaerella*, *Phyllactinia* such as *Phyllactinia kaki*, *Gloesporium* such as *Gloesporium kaki*, *Gymnosporangium* such as *Gymnosporangium yamadae*, *Leptotthrydium* such as *Leptotthrydium pomi*, *Podosphaera* such as *Podosphaera leucotricha*, *Gloedes* such as *Gloedes pomigena*, *Cladosporium* such as *Cladosporium carpophilum*, *Phomopsis*, *Phytopora*, *Phytophthora* such as *Phytophthora infestans*, *Verticillium*, *Glomerella* such as *Glomerella cingulata*;

## 5

Drechslera; Bipolaris; Personospora; Phaeoisariopsis such as Phaeoisariopsis vitis; Spaceloma such as Spaceloma ampelina; Pseudocercospora such as Pseudocercospora herpotrichoides; Pseudoperonospora; Puccinia; Typhula; Pyricularia such as Pyricularia oryzae; Rhizoctonia; Stachosporium such as Stachosporium nodorum; Uncinula such as Uncinula necator; Ustilago; Gaeumannomyces species such as Gaeumannomyces graminis and Fusarium such as Fusarium dimerium, Fusarium merismoides, Fusarium lateritium, Fusarium decemcellulare, Fusarium poae, Fusarium tricinctum, Fusarium sporotrichioides, Fusarium chlamydosporum, Fusarium moniliforme, Fusarium proliferatum, Fusarium anthophilum, Fusarium subglutinans, Fusarium hygumai, Fusarium oxysporum, Fusarium solani, Fusarium culmorum, Fusarium sambucinum, Fusarium crookwellense, Fusarium avenaceum ssp. avenaceum, Fusarium avenaceum ssp. aywarte, Fusarium avenaceum ssp. nurragi, Fusarium heterosporum, Fusarium acuminatum ssp. acuminatum, Fusarium acuminatum ssp. armeniacum, Fusarium longipes, Fusarium compactum, Fusarium equiseti, Fusarium scripi, Fusarium polyphialidicum, Fusarium semitectum and Fusarium beomiforme and especially preferred from the genera Fusarium such as Fusarium graminearum, most preferred from the group consisting of the genera and species Fusarium, Fusarium dimerium, Fusarium merismoides, Fusarium lateritium, Fusarium decemcellulare, Fusarium poae, Fusarium tricinctum, Fusarium sporotrichioides, Fusarium chlamydosporum, Fusarium moniliforme, Fusarium proliferatum, Fusarium anthophilum, Fusarium subglutinans, Fusarium nygamai, Fusarium oxysporum, Fusarium solani, Fusarium culmorum, Fusarium sambucinum, Fusarium crookwellense, Fusarium avenaceum ssp. avenaceum, Fusarium avenaceum ssp. aywarte, Fusarium avenaceum ssp. nurragi, Fusarium heterosporum, Fusarium acuminatum ssp. acuminatum, Fusarium acuminatum ssp. armeniacum, Fusarium longipes, Fusarium compactum, Fusarium equiseti, Fusarium scripi, Fusarium polyphialidicum, Fusarium semitectum and Fusarium beomiforme wherein Fusarium graminearum is most preferred.

35

The host organism in which the origin of replication a) is functionally active essentially serves for constructing and propagating the plasmid vector of the invention. The host organism must be genetically different from the filamentous fungi to be transformed, since replication of the plasmid vector should not take place in the filamentous fungi to be transformed but is desired in the host organism, due to using the origin of replication a). Host organisms which may be used are all common microorganisms which can easily be manipulated by genetic engineering. Preferred host organisms are Gram-negative bacteria such as the genera Escherichia and Salmonella e.g. Escherichia coli and Salmonella typhimurium or Gram-positive bacteria such as the genera Bacil-

45

lus and Streptomyces, e.g. Bacillus subtilis and Streptomyces nidulans. Particularly preferred are Gram-negative bacteria such as Escherichia, e.g. Escherichia coli.

- 5 Preferred origins of replications (ori) are the col E1 ori, the fl ori.

The term "selection marker for a host organism" set forth in b) means a gene or the expression product of the gene. Preferred  
10 meanings are genes whose expression causes resistance of the host organism to antibiotics, by preference a resistance to kanamycin, chloramphenicol, tetracycline, zeocin or ampicillin and particularly preferred ampicillin and kanamycin.

- 15 In a preferred embodiment, the element a) of the plasmid vector according to the invention comprises a col E1 origin of replication and the ampicillin resistance gene as selection marker for the host organism.

- 20 The element c) is hereinbelow termed as "hygromycin cassette". The coding region of the hygromycin resistance gene (hereinbelow termed "hygromycin gene") is known by the skilled artisan (Gritz L. and Davies J., 1983, Gene 25, 179-188, Kaster, K.R., Burgett S.G. and Ingolia T.D., 1984, Curr. Genet. 8, 353-358) and has a  
25 length of 1026bp.

Examples of suitable promoters to which the coding region of the hygromycin gene is functionally linked to, are the GPD-1-, PX6-, TEF-, CUP1-, PGK-, GAP1-, TPI, PHO5-, AOX1, GAL10/CYC-1, CYC1,  
30 OliC-, ADH-, TDH-, Kex2-, MF $\alpha$ -, or the NMT-promotor (Degryse et al., Yeast 1995 Jun 15; 11(7):629-40; Romanos et al. Yeast 1992 Jun;8(6):423-88; Benito et al. Eur. J. Plant Pathol. 104, 207-220 (1998); Cregg et al. Biotechnology (N Y) 1993 Aug;11(8):905-10; Luo X., Gene 1995 Sep 22;163(1):127-31; Nacken et al., Gene 1996  
35 Oct 10;175(1-2):253-60; Turgeon et al., Mol Cell Biol 1987 Sep;7(9):3297-305), preferably the CYC1-, ADH-, TDH-, Kex2-, GPD-1-, PX6, TEF-, CUP1-, PGK-, GAP1-, TPI, PHO5- or AOX1-promotor, more preferably the GPD-1-, PX6, TEF- or the CUP1-promotor, most preferably the GPD1 or the TEF-promotor.

40 Examples of suitable terminators that are functionally linked to the coding region of the hygromycin gene are the AOX1-, nos-, PGK-, TrpC- or the CYC1-terminator (Degryse et al., Yeast 1995 Jun 15; 11(7):629-40; Brunelli et al. Yeast 1993 Dec9(12):  
45 1309-18; Frisch et al., Plant Mol. Biol. 27 (2), 405-409 (1995); Scorer et al., Biotechnology (N.Y.) 12 (2), 181-184 (1994), Genbank acc. number 246232; Punt et al., (1987) Gene 56 (1),

117-124)), preferably the CYC1- or nos-terminator, more preferably the nos-terminator.

In a preferred embodiment, the hygromycin cassette comprises a 5 GPD-1 promotor functionally linked to the coding region of the hygromycin region which is functionally linked to the nos-terminator.

A functional linkage is understood as meaning the sequential arrangement of promoter and coding sequence, of coding sequence and terminator or of promoter, coding sequence and terminator in such a manner that each of the regulatory elements can, upon expression of the coding sequence, fulfil its function upon the recombinant expression of the nucleic acid sequence. Direct linkage in the chemical sense is not necessarily required for this purpose. Preferred arrangements are those in which the hygromycin gene to be expressed recombinantly is positioned downstream of the sequence which acts as promoter, so that the two sequences are linked covalently to each other. The distance between the promoter sequence and the nucleic acid sequence to be expressed recombinantly is preferably less than 100 base pairs, especially preferably less than 50 base pairs, very especially preferably less than 10 base pairs. The distance between the terminator sequence and the nucleic acid sequence to be expressed recombinantly is preferably less than 100 base pairs, especially preferably less than 50 base pairs, very especially preferably less than 10 base pairs. However, further sequences which, for example, exert the function of a linker with certain restriction enzyme cleavage sites, or of a signal peptide, may also be positioned between the two sequences.

These vectors are not only much more smaller than currently used plasmid vectors, but exhibit also a high transformation efficiency. Surprisingly, a high transformation efficiency can be gained even if small DNA-fragments of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp of the nucleic acid sequence d) to be analyzed are used. The average degree of illegitimate recombination is below 30%, preferably below 25%, more preferably 20%, most preferably between 0 to 15%.

The nucleic acid sequence d) has a homology of at least 80% to the nucleic acid sequence of the filamentous fungi to be transformed, preferably at least 90%, more preferably at least 95% and most preferably at least 100%.

8

In a preferred embodiment, the nucleic acid sequence d) originates from a filamentous fungi and has a length of at least 300bp, preferably 400bp, more preferably at least 450bp, most preferably at least 500bp. These lengths are suitable for functional genomic studies for which a high number of transformants is required. Also nucleic acid sequences exceeding 500bp can be used, e.g. for the purpose recombinant expression.

If the nucleic acid sequence d) is to be expressed recombinantly in the filamentous fungi, it can be functionally linked to a promoter e) and optionally to a terminator f).

Examples of suitable promoters e) are for example the AUG1-, GPD1-, PX6-, TEF-, CUP1-, PGK-, GAP1-, TPI, PHO5-, AOX1-, GAL10/CYC-1, CYC1, OliC-, ADH-, TDH-, Kex2-, MF $\alpha$ - or the NMT-promotor or combinations of the aforementioned promoters (Degryse et al., Yeast 1995 Jun 15; 11(7):629-40; Romanos et al. Yeast 1992 Jun;8(6):423-88; Benito et al. Eur. J. Plant Pathol. 104, 207-220 (1998); Cregg et al. Biotechnology (N Y) 1993 Aug;11(8):905-10; Luo X., Gene 1995 Sep 22;163(1):127-31; Nacken et al., Gene 1996 Oct 10;175(1-2): 253-60; Turgeon et al., Mol Cell Biol 1987 Sep;7(9):3297-305).

Examples of suitable terminators f) are the NMT-, Gcyl-, TrpC-, AOX1-, nos-, the PGK- or the CYC1-terminator (Degryse et al., Yeast 1995 Jun 15; 11(7):629-40; Brunelli et al. Yeast 1993 Dec9(12): 1309-18; Frisch et al., Plant Mol. Biol. 27 (2), 405-409 (1995); Scorer et al., Biotechnology (N.Y.) 12 (2), 181-184 (1994), Genbank acc. number Z46232; Zhao et al. Genbank acc number : AF049064; Punt et al., (1987) Gene 56 (1), 117-124).

The nucleic acid sequence d) can be also functionally linked to an affinity tag to purify the encoded protein and/or to a reporter gene to study biochemical properties of the nucleic acid sequence d) in vivo, respectively.

"Reporter genes" encode readily quantifiable proteins. Using these genes, an assessment of transformation efficacy or of the site or time of expression can be made via growth, fluorescence, chemoluminescence, bioluminescence or resistance assay or via photometric measurement (intrinsic color) or enzyme activity. Very especially preferred in this context are reporter proteins (Schenborn E, Groskreutz D. Mol. Biotechnol. 1999; 13(1):29-44) such as the "green fluorescence protein" (GFP) (Gerdes HH and Kaether C, FEBS Lett. 1996; 389(1):44-47; Chui WL et al., Curr. Biol. 1996, 6:325-330; Leffel SM et al., Biotechniques.

23(5):912-8, 1997), chloramphenicol acetyl transferase, a luciferase (Giacomin, Plant Sci. 1996, 116:59-72; Scikantha, J. Bact. 1996, 178:121; Millar et al., Plant Mol. Biol. Rep. 1992 10:324-414), and luciferase genes, in general  $\beta$ -galactosidase or  
5  $\beta$ -glucuronidase (Jefferson et al., EMBO J. 1987, 6, 3901-3907), the Ura3 gene, the Ilv2 gene, the 2-desoxyglucose-6-phosphate phosphatase gene,  $\beta$ -lactamase gene, the neomycin phosphotransferase gene, the hygromycin phosphotransferase gene, or the BASTA (= gluphosinate) resistance gene.

10

The term "affinity tag" denotes a peptide or polypeptide whose coding nucleic acid sequence can be fused to the nucleic acid sequence d) either directly or using a linker, by customary cloning techniques. The affinity tag serves to isolate the recombinant  
15 target protein by means of affinity chromatography. The abovementioned linker can optionally comprise a protease cleavage site (for example for thrombin or factor Xa), whereby the affinity tag can be cleaved off from the target protein, as required. Examples of customary affinity tags are the "his-tag", for example from  
20 Quiagen, Hilden, "strep-tag", "myc-tag" (Invitrogen, Carlsberg), New England Biolab's tag which consists of a chitin binding domain and an intein, and what is known as the CBD-tag from Novagen.

25 In a particularly preferred embodiment, the plasmid vector comprises an coli E1 ori, the ampicillin resistance gene as selection marker, a GPD-1 promotor functionally linked to the coding region of the hygromycin resistance gene which is functionally linked to the nos-terminator.

30

Preferably, the vector also comprises a multiple cloning site comprising appropriate restriction enzyme site. Appropriate restriction sites are well known by the skilled artisan.

35 In a further preferred embodiment, the plasmid vector additionally comprises a TA-cloning site to facilitate the overall cloning procedure.

40 Examples of particularly preferred embodiments are set forth in Fig. 1 and 2.

All of the above-mentioned embodiments of plasmid vectors are hereinbelow termed as "plasmid vector (or vector) according to the invention".

45

## 10

A vector according to the invention may also comprise at least one additional selection marker.

If a plasmid is used for recombinant expression in a host, a marker is required indicating the successful transfer of the plasmid vector DNA into the filamentous fungi to be transformed.

Surprisingly, we have found that the gene fragments of the polyketide synthase are a well suited selection marker. The term "selection marker" referred to the polyketide synthase herein means a nucleic acid sequence.

More precisely, the term "selectable marker", "selection marker" or "marker" used in connection with polyketide synthetase for transformation of filamentous fungi means a nucleic acid sequence encoding a polyketide synthetase or fragments of the aforementioned nucleic acid sequence. Preferred embodiments of the aforementioned marker as well as preferred embodiments of methods of use of the respective marker are described herein below.

Polyketide synthases are multifunctional enzymes that are involved in the biosynthesis of several important polyketides. Polyketides constitute a large and highly diverse group of secondary metabolites, synthesized by bacteria, fungi and plants and algae. They include antibiotics, compounds with mycotoxic activity, and compounds within pigment biosynthetic pathways. Further a polyketide synthase is described to be required for fungal virulence of *Cochliobolus heterostrophus* toward maize (Yang et al., 1996 PMID:8953776). Polyketide Synthetases are furthermore known from *Wangiella dermatidis* (PubMedID:11179356), from *Aspergillus nidulans* (Swiss-prot ID: Q03149) and from *Aspergillus parasiticus* (Swiss-Prot ID:Q12053).

The use of polyketide synthase as selectable marker to be used in an expression for filamentous fungi has not yet been described.

The present invention also encompasses a selection marker comprising a nucleic acid sequence encoding a polyketide synthetase fragment, wherein said nucleic acid sequence comprises

- i. a nucleic acid sequence shown in SEQ ID NO:1 or SEQ ID NO:2; or
- ii. parts of the nucleic acid sequence as defined in i. consisting of at least 300bp.

## 11

Parts or segments of nucleic acid sequences set forth in ii. consist of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp of the nucleic acid sequences. In a further preferred embodiment, those 5 parts are selected from SEQ ID NO:1, preferably from 732bp to 5881bp of SEQ ID NO:1 e.g. from 2236bp to 2870bp.

Furthermore, the present invention encompasses a selection marker comprising a nucleic acid sequence encoding a polyketide synthetase or a polyketide synthetase fragment, wherein said nucleic acid sequence comprises

- i. a nucleic acid sequence shown in SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5; or
- ii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or
- iii. a functional equivalent of the nucleic acid sequence set forth in i) which is encoded by an amino acid sequence that has at least an identity of 50% with the SEQ ID NO:6; or
- iv. parts of the nucleic acid sequence as defined in i., ii. or iii. consisting of at least 300bp.
- v. parts of the nucleic acid sequence as defined in i., ii. or iii. consisting of at least 300bp comprising
  - a) a nucleic acid sequence shown in SEQ ID NO:7; or
  - b) a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:8 by back translation; or
  - c) a functional equivalent of a nucleic acid sequence set forth in a) which is encoded by an amino acid sequence that has at least an identity of 85% with the SEQ ID NO:8.

Parts or segments of nucleic acid sequences set forth in iii. or v. consist of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp of the nucleic acid sequences. Preferably, the aforementioned parts or segments of nucleic acid sequences are those set forth in v.a), v.b) or v.c), more preferably those set forth in v.a) or v.b) most preferably those set forth in v.a). For example, those parts

12

can be selected from 2234bp to 2865bp of SEQ ID NO:3.

The functional equivalents of the nucleic acid sequence set forth in iv. are encoded by an amino acid sequence that has at least an identity of 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64% or 65% or preferred of 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% or 79% more preferred of 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% most preferred of 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% with SEQ ID NO:6.

The functional equivalents of the nucleic acid sequence set forth in v.c) are encoded by an amino acid sequence that has at least an identity of 85%, 86%, 87% or 88% or preferred of 89%, 90%, 91%, 92% or 93% more preferred of 94%, 95% or 96% most preferred of 97%, 98% or 99% with SEQ ID NO:8.

Preferred are nucleic acid sequences as defined above originate from filamentous fungi, preferably phytopathogenic filamentous fungi selected from the group consisting of the genera *Neurospora*, *Alternaria*, *Podosphaera*, *Sclerotinia*, *Physalospora*, *Botrytis*, *Corynespora*, *Colletotrichum*, *Diplocarpon*, *Elsinoe*, *Diaporthe*, *Sphaerotheca*, *Cinula*, *Cercospora*, *Erysiphe*, *Sphaerotheca*, *Leveillula*, *Mycosphaerella*, *Phyllactinia*, *Gloesporium*, *Gymnosporangium*, *Leptotthrydium*, *Podosphaera*, *Gloeodes*, *Cladosporium*, *Phomopsis*, *Phytopora*, *Phytophthora*, *Erysiphe*, *Fusarium*, *Verticillium*, *Glomerella*, *Drechslera*, *Bipolaris*, *Personospora*, *Phaeoisariopsis*, *Spaceloma*, *Pseudocercospora*, *Pseudoperonospora*, *Puccinia*, *Typhula*, *Pyricularia*, *Rhizoctonia*, *Stachosporium*, *Uncinula*, *Ustilago*, *Gaeumannomyces* and *Fusarium*, more preferred from the group consisting of the genera and species *Alternaria*, *Podosphaera*, *Sclerotinia*, *Physalospora* such as *Physalospora* canker, *Botrytis* species such as *Botrytis cinerea*, *Corynespora* such as *Corynespora melonis*, *Colletotrichum*, *Diplocarpon* such as *Diplocarpon rosae*, *Elsinoe* such as *Elsinoe fawcetti*, *Diaporthe* such as *Diaporthe citri*, *Sphaerotheca*, *Cinula* such as *Cinula neccata*, *Cercospora*, *Erysiphe* such as *Erysiphe cichoracearum* and *Erysiphe graminis*, *Sphaerotheca* such as *Sphaerotheca fuliginea*, *Leveillula* such as *Leveillula taurica*, *Mycosphaerella*, *Phyllactinia* such as *Phyllactinia kakicola*, *Gloesporium* such as *Gloesporium kaki*, *Gymnosporangium* such as *Gymnosporangium yamadae*, *Leptotthrydium* such as *Leptotthrydium pomi*, *Podosphaera* such as *Podosphaera leucotricha*, *Gloeodes* such as *Gloeodes pomigena*, *Cladosporium* such as *Cladosporium carpophilum*, *Phomopsis*, *Phytopora*, *Phytophthora* such as *Phytophthora infestans*, *Verticillium*, *Glomerella* such as *Glomerella cingulata*, *Drechslera*, *Bipolaris*, *Personospora*, *Phaeoisariopsis* such as *Phaeoisariopsis vitis*, *Spaceloma* such as *Spaceloma ampe-*

- lina; Pseudocercospora such as Pseudocercospora herpotrichoides; Pseudoperonospora; Puccinia; Typhula; Pyricularia such as Pyricularia oryzae; Rhizoctonia; Stachosporium such as Stachosporium nodorum; Uncinula such as Uncinula necator; Ustilago;
- 5 Gaeumannomyces species such as Gaeumannomyces graminis and Fusarium such as Fusarium dimerium, Fusarium merismoides, Fusarium lateritium, Fusarium decemcellulare, Fusarium poae, Fusarium tricinctum, Fusarium sporotrichioides, Fusarium chlamydosporum, Fusarium moniliforme, Fusarium proliferatum, Fusarium anthophilum,
- 10 Fusarium subglutinans, Fusarium nygamai, Fusarium oxysporum, Fusarium solani, Fusarium culmorum, Fusarium sambucinum, Fusarium crookwellense, Fusarium avenaceum ssp. avenaceum, Fusarium avenaceum ssp. aywerte, Fusarium avenaceum ssp. nurragi, Fusarium heterosporum, Fusarium acuminatum ssp. acuminatum, Fusarium acuminatum ssp. armeniacum, Fusarium longipes, Fusarium compactum, Fusarium equiseti, Fusarium scripi, Fusarium polyphialidicum, Fusarium semitectum and Fusarium beomiforme and especially preferred from the genera Fusarium such as Fusarium graminearum, most preferred from the group consisting of the genera and species Fusarium,
- 20 Fusarium dimerium, Fusarium merismoides, Fusarium lateritium, Fusarium decemcellulare, Fusarium poae, Fusarium tricinctum, Fusarium sporotrichioides, Fusarium chlamydosporum, Fusarium moniliforme, Fusarium proliferatum, Fusarium anthophilum, Fusarium subglutinans, Fusarium nygamai, Fusarium oxysporum, Fusarium solani, Fusarium culmorum, Fusarium sambucinum, Fusarium crookwellense, Fusarium avenaceum ssp. avenaceum, Fusarium avenaceum ssp. aywerte, Fusarium avenaceum ssp. nurragi, Fusarium heterosporum, Fusarium acuminatum ssp. acuminatum, Fusarium acuminatum ssp. armeniacum, Fusarium longipes, Fusarium compactum, Fusarium equiseti, Fusarium scripi, Fusarium polyphialidicum, Fusarium semitectum and Fusarium beomiforme wherein Fusarium graminearum is most preferred.
- 30

- Preferred non-phytopathogenic filamentous fungi are fungi of
- 35 group consisting of the genera Neurospora such as Neurospora crassa, Aspergillus such as Aspergillus parasiticus, Aspergillus nidulans, Aspergillus niger and Wangiella such as Wangiella dermatidis.
- 40 The term "comprising" means that the nucleic acid sequence according to the invention can be flanked by additional nucleic acid sequences that have on the 5' end a sequence length of at least 1000 bp and preferably at least 500 bp, more preferably at least 100bp, most preferably at least 50bp and on the 3' a sequence length of at least 1000 bp and preferably at least 500 bp,
- 45 more preferably at least 100 bp most preferably at least 50bp.

## 14

"Functional equivalents" in the present context describe nucleic acid sequences which hybridize under standard conditions with the nucleic acid sequence or portions of the nucleic acid sequence having the function of the a selection marker.

5

It is advantageous to use short oligonucleotides of a length of 10-50 bp, preferably 15-40 bp, for example of the conserved or other regions, which can be determined via comparisons with other related genes in a manner known to the skilled worker for the hybridization. Alternatively, it is also possible to use longer fragments of the nucleic acids according to the invention or the complete sequences for the hybridization. These standard conditions vary depending on the nucleic acid used, viz. oligonucleotide, longer fragment or complete sequence, or depending on which type of nucleic acid, viz. DNA or RNA, is being used for the hybridization. Thus, for example, the melting temperatures for DNA:DNA hybrids are approx. 10°C lower than those of DNA:RNA hybrids of equal length.

20

Standard conditions are understood as meaning, depending on the nucleic acid, for example temperatures between 42 and 58°C in an aqueous buffer solution with a concentration of between 0.1 and 5 x SSC (1 x SSC = 0.15 M NaCl, 15 mM sodium citrate, pH 7.2) or additionally in the presence of 50% formamide such as, for example, 42°C in 5 x SSC, 50% formamide. The hybridization conditions for DNA:DNA hybrids are advantageously 0.1 x SSC and temperatures of between approximately 20°C and 45°C, preferably between approximately 30°C and 45°C. The hybridization conditions for DNA:RNA hybrids are advantageously 0.1 x SSC and temperatures of between approximately 30°C and 55°C, preferably between approximately 45°C and 55°C. These temperatures stated for the hybridization are melting temperature values which have been calculated by way of example for a nucleic acid with a length of approx. 100 nucleotides and a G + C content of 50% in the absence of formamide. The experimental conditions for DNA hybridization are described in specialist textbooks of genetics such as, for example, Sambrook et al., "Molecular Cloning", Cold Spring Harbor Laboratory, 1989 and can be calculated using formulae known to the skilled worker, for example as a function of the length of the nucleic acids, the type of the hybrids or the G + C content. The skilled worker can find more information on hybridization in the following textbooks: Ausubel et al. (eds), 1985, Current Protocols in Molecular Biology, John Wiley & Sons, New York; Hames and Higgins (eds), 1985, Nucleic Acids Hybridization: A Practical Approach, IRL Press at Oxford University Press, Oxford; Brown (ed), 1991, Essential Molecular Biology: A Practical Approach, IRL Press at Ox-

45

ford University Press, Oxford.

A functional equivalent is furthermore also understood as meaning, in particular, natural or artificial mutations of the relevant nucleic acid sequences of the polyketide synthetase (PKS) as set forth in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5 and its homologs from other organisms, wherein mutations comprise substitutions, additions, deletions, inversions or insertions of one or more nucleotide residues. This may also lead to a modification of the corresponding amino acid sequence of the PKS by substitution, insertion or deletion of one or more amino acids.

Thus, the scope of the present invention also extends to, for example, those nucleotide sequences which are obtained by modification of the nucleic acid sequence of the selection marker described by SEQ ID NO:1 or by SEQ ID NO:2 or SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5 respectively. The purpose of such a modification can be, for example, the insertion of further cleavage sites for restriction enzymes, the removal of excess DNA, or the addition of further sequences. Said nucleic acid sequences should still maintain the desired function as marker for targeted transformation, despite the deviating nucleic acid sequence.

The term "identity" or "homology" between two nucleic acid sequences or polypeptide sequences is defined by the identity of the nucleic acid sequence/polypeptide sequence by in each case the entire sequence length, which is calculated by alignment with the aid of the program algorithm GAP (Wisconsin Package Version 10.0, University of Wisconsin, Genetics Computer Group (GCG), Madison, USA), setting the following parameters:

Gap Weight: 8

Length Weight: 4

Average Match: 2,912

Average Mismatch:-2,003

The term homology if used herein is the same as the term identity.

Functional equivalents thus encompass naturally occurring variants of the sequences described herein, and also artificial, for example chemically synthesized, nucleic acid sequences adapted to the codon usage, or the amino acid sequences derived therefrom.

## 16

Moreover, SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5, nucleic acid sequences derived from the amino acid sequence SEQ ID NO:6 by back translation or parts of the aforementioned nucleic acid sequences can be used for the detection and isolation of functional equivalents of from other fungi on the basis of sequence identities. In this context, part or all of the sequence of the SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5 or nucleic acid sequences derived from the amino acid sequence SEQ ID NO:6 by back translation can be used as probe (e.g. hybridization probe) for screening in a genomic library or a cDNA library of the fungal species in question or in a computer search for sequences of functional equivalents in electronic databases. Especially for computer search for sequences of functional equivalents in electronic databases, the amino acid sequence SEQ ID NO:6 or parts of the amino acid sequence SEQ ID NO:6 are usefull.

For the preparation of hybridization probes, SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5 or parts of the aforementioned nucleic acid sequences can be used. The preparation of these probes and the experimental procedure are known. For example, this can be effected via the tailor-made preparation of radioactive or nonradioactive probes by means of PCR and the use of suitably labeled oligonucleotides, followed by hybridization experiments. The technologies required for this purpose are given, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1989). The probes in question can furthermore be modified by standard technology (lit. SDM or random mutagenesis) in such a way that they can be employed for other purposes, for example as probe which hybridizes specifically with mRNA and the corresponding coding sequences in order to analyze the corresponding sequences in other organisms.

Furthermore, the cDNA could be used to engineer recombinant microorganisms to produce polyketide agents of pharmaceutical or agricultural interest as described by Pfeifer et al. (Pfeifer BA, Admiraal SJ, Gramajo H, Cane DE, Khosla C., Science 2001 Mar 2;291(5509):1790-2). Thus, the present invention also comprises polypeptides with the biological activity of a polyketide synthetase encoded by an nucleic acid sequence comprising

i. a nucleic acid sequence shown in SEQ ID NO:5 or

ii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or

5 iii. nucleic acid sequence which is encoded by a functional analogue of an amino acid sequence that has at least an identity of 50% with the SEQ ID NO:6.

The term "functional analogues" describes nucleic acid sequences which are capable of bringing about the expression, in a filamentous fungi, of a polypeptide with the biological activity of polyketide synthetase and which can be deduced from an amino acid sequence by back translation which has a defined degree of identity with SEQ ID NO:6. The functional analogues set forth in iii) have at least an identity of 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64% or 65% or preferred of 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% or 79% more preferred of 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% most preferred of 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% with the SEQ ID NO:6.

Thus, the present invention also encompasses, for example, those nucleotide sequences which are obtained by modification of the abovementioned nucleic acid sequences. For example, such modifications can be generated by techniques with which the skilled worker is familiar, such as "site directed mutagenesis", "error prone PCR", "DNA shuffling" (Nature 370, 1994, pp.389-391) or "staggered extension process" (Nature Biotechnol. 16, 1998, pp.258-261). The purpose of such a modification can be, for example, the insertion of further cleavage sites for restriction enzymes, the removal of DNA in order to truncate the sequence, the substitution of nucleotides in order to optimize the codons, or the addition of further sequences. Proteins which are encoded via modified nucleic acid sequences must retain the desired functions despite a deviating nucleic acid sequence.

Functional analogues thus comprise naturally occurring variants of the herein-described sequences and artificial nucleic acid sequences, for example those which have been obtained by chemical synthesis and which are adapted to the codon usage, and also the amino acid sequences derived from them.

As explained above, also the expression cassette or the vector comprising a PKS encoding nucleic acid sequence may comprise at least an additional selection marker, preferably the hygromycin resistance gene so that in a particular preferred embodiment, the selection of the successfully transformed filamentous fungi can be

18

carried out by hygromycin resistance of successfully transformed clones and by the presence of pigment (colour) of successfully transformed clones. Most preferably, the vector comprising the PKS encoding nucleic acid sequence is a vector according to the invention comprising a PKS encoding nucleic acid sequence. In addition to the aforementioned selection methods homologous recombination can be confirmed by PCR based on oligonucleotides preferably derived from the vector sequence flanking the 5 and 3 region of the gene to be inserted. Specific examples of these primers are given in the examples.

The invention furthermore relates to the use of polyketide synthetase encoding nucleic acid sequences as marker for targeted transformation in filamentous fungi.

15

Preferably, the present invention comprises the use of a nucleic acid sequence comprising

20

- a) a nucleic acid sequence encoding a polyketide synthetase; or
- b) parts of the nucleic acid sequence as defined in i., ii. or iii. consisting of at least 300bp

for transformation of filamentous fungi.

25

Preferred is the use of a nucleic acid sequence as marker for targeted transformation in filamentous fungi said nucleic acid comprising

- 30 i. a nucleic acid sequence shown in SEQ ID NO:1 or SEQ ID NO:2; or

- ii. parts of the nucleic acid sequence as defined in i. consisting of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp;

35

Equally preferred is the use of a nucleic acid sequence as marker for targeted transformation in filamentous fungi said nucleic acid comprising

40

- iii. a nucleic acid sequence shown in SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5; or

45

- iv. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or

## 19

v. a functional equivalent of the nucleic acid sequence set forth in i) or iii) which is encoded by an amino acid sequence that has at least an identity of 40% with the SEQ ID NO:6; or

5

vi. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO:11;

10

vii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation; or

15

viii. parts of the nucleic acid sequence as defined in iii., iv, v., vi or vii. consisting of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp; or

20

ix. parts of the nucleic acid sequence as defined in iii., iv, v., vi or vii. consisting of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp comprising

25

a) a nucleic acid sequence shown in SEQ ID NO:7; or

b) a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:8 by back translation; or

30

c) a functional equivalent of a nucleic acid sequence set forth in a) which is encoded by an amino acid sequence that has at least an identity of 68% with the SEQ ID NO:8.

35 The nucleic acid sequences according to i. to ix encode for a polypeptid with the biological function of a polyketide synthetase or for a fragment of the aforementioned polypeptide.

Under the aforementioned sequences, the nucleic acid sequences according to i., ii., iii., iv., v. as well as parts of the aforementioned nucleic acid sequence consisting of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp are preferred. Those parts are preferably those set forth in ix.

45

Preferred phytopathogenic and non-phytopathogenic filamentous fungi are those mentioned above. The aforementioned nucleic acid sequences are hereinbelow also termed "PKS-marker". Preferably, the term "PKS-marker" designates nucleic acid sequences according to i., ii., iii., iv., v. as well as parts of the aforementioned nucleic acid sequence consisting of at least 300bp, preferably at least 400bp, more preferably at least 450bp, most preferably at least 500bp are preferred.

10 The functional equivalents of the nucleic acid sequence set forth in iv. can be deduced from a functional equivalent of the amino acid sequence shown in SEQ ID NO:6 by back translation having at least an identity of 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48% or 49% preferred of 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% or 79% more preferred of 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% most preferred of 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% with the SEQ ID NO:6:

20

The functional equivalents of the nucleic acid sequence set forth in ix.c) can be deduced from a functional equivalent of the amino acid sequence shown in SEQ ID NO:8 by back translation having at least an identity of 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77% or 78% preferred of 79%, 80%, 81%, 82%, 83%, 84% or 85% more preferred of 86%, 87%, 88%, 89% or 90% most preferred of 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% with the SEQ ID NO:8.

The use PKS marker for targeted transformation of filamentous fungi can be based on significant reduction in the amount polyketide synthetase, which is present in a filamentous fungi. A reduction in the amount the polyketide synthetase means that the amount of polypeptide is reduced via recombinant methods. Preferred is a reduction of at least 30%, more preferred of at least 50%, most preferred by at least 50%, up to 100% reduction (blocking) relative to the amount of polyketide synthetase present in the respective wild-type.

Reduction via recombinant methods can involve "antisense techniques", which describes a technology for the suppression (reduction) of expression of polyketide synthetase, where a PKS-marker is transformed into the respective filamentous fungi in "antisense" orientation under the control of a suitable promoter. This method is used preferably for *Aspergillus* species, more preferably for *Aspergillus nidulans*. The technologies required herefore are well known by the skilled artisan (for example see Bautista et al., Appl. Environ. Microbiol. 2000; 66(10) 4579-81).

Suitable vectors therefore comprise an expression cassette comprising

1. An expression cassette comprising

5

a) a promotor sequence in functional linkage with a PKS-marker in antisense orientation; and optionally

10

b) further genetic control sequences in functionally linked with a nucleic acid sequence according to a).

The afore-mentioned expression cassette is hereinbelow termed as "PKS-Marker-expression cassette".

15 The term "expression cassette" can be defined as follows: An expression cassette comprises a nucleic acid sequence, which should be expressed, linked functionally to at least one genetic control element, such as a promoter, and, advantageously, a further control element, such as a terminator. Examples of suitable promoters and terminators are given above. The nucleic acid sequence of the expression cassette can be, for example, a genomic or complementary DNA sequence or an RNA sequence, and the semisynthetic or fully synthetic analogs thereof. These sequences can exist in linear or circular form, extrachromosomally or integrated into the genome. The nucleic acid sequences in question can be synthesized or obtained naturally or comprise a mixture of synthetic and natural DNA components, and consist of a variety of heterologous gene segments from various organisms.

30 Artificial nucleic acid sequences are also suitable in this context as long as they make possible the expression, in a cell or organism, of a polypeptide encoded by a nucleic acid sequence according to the invention and having the biological activity of a polyketide synthetase. For example, synthetic nucleotide sequences can be generated which have been optimized with regard to the codon usage of the organisms to be transformed.

All of the abovementioned nucleotide sequences can be generated from the nucleotide units by chemical synthesis in the manner known per se, for example by fragment condensation of individual, overlapping complementary nucleotide units of the double helix. Oligonucleotides can be synthesized chemically for example in the manner known per se using the phosphoamidite method (Voet, Voet, 2nd Edition, Wiley Press New York, pp. 896-897). When preparing an expression cassette, various DNA fragments can be manipulated in such a way that a nucleotide sequence with the correct direction of reading and the correct reading frame is obtained. The

## 22

nucleic acid fragments are linked to each other via general cloning techniques as are described, for example, in T. Maniatis, E.F. Fritsch and J. Sambrook, "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 5 (1989) and in T.J. Silhavy, M.L. Berman and L.W. Enquist, Experiments with Gene Fusions, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY (1984) and in Ausubel, F.M. et al., "Current Protocols in Molecular Biology", Greene Publishing Assoc. and Wiley-Interscience (1994).

10

The term "genetic control element" describes sequences which have an effect on the transcription and, if appropriate, translation of the nucleic acids according to the invention in prokaryotic or eukaryotic organisms. Examples are terminators. Examples of suitable 15 terminators are given above. In addition to the aforementioned control sequences, or instead of these sequences, the natural regulation of these sequences may still be present before the actual structural genes and may, if appropriate, have been modified genetically in such a way that the natural regulation 20 has been switched off and the expression of the target gene has been modified, that is to say increased or reduced. The choice of the control sequence depends on the host organism or starting organism. Genetic control sequences furthermore also comprise the 5'-untranslated region, introns or the noncoding 3' region of 25 genes. Control sequences are furthermore understood as meaning those which make possible a homologous recombination or insertion into the genome of a host organism or which permit the removal from the genome. Genetic control sequences also comprise further promoters, promoter elements or minimal promoters.

30

The transcription of the PKS marker leads to suppression of the transcription of the natural polyketide synthetase gene, which can be detected by loss of colour of the transformed fungi relative to the respective wild-type strain.

35

In a preferred embodiment, the reduction via recombinant methods is based on a gene knock out of the polyketide synthetase gene using either an expression cassette additionally comprising the PKS-marker or a vector comprising the PKS marker in the respective filamentous fungi. Disruption of the PKS marker will lead to 40 a loss of colour.

Preferred phytopathogenic filamentous fungi are those mentioned above. Preferred non-phytopathogenic filamentous fungi are fungi 45 of group consisting of the genera Aspergillus such as Aspergillus

## 23

parasiticus, Aspergillus nidulans and Wangiella such as Wangiella dermatidis.

In this connection, the selection of the functional equivalent  
5 for the use as marker gene depends on the fungi to be transformed. By preference, the polyketide synthetase fragment has an identity of at least 80%, by preference at least 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, especially preferably at least 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% identity with the  
10 polyketide synthetase of the fungi to be transformed.

For example, for transformation of Fusarium graminearum, a nucleic acid sequence can be selected comprising a nucleic acid sequence comprising

15

i. a nucleic acid sequence shown in SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5; or

ii. a nucleic acid sequence that has at least an identity of 80%  
20 SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5; or

iii. parts of the nucleic acid sequence as defined in i. or ii.  
consisting of at least 300bp, preferably at least 400bp,  
25 more preferably at least 450bp, most preferably at least 500bp.

iv. parts of the nucleic acid sequence as defined in i. or ii.  
consisting of at least 300bp, preferably at least 400bp,  
30 more preferably at least 450bp, most preferably at least 500bp comprising

a) a nucleic acid sequence shown in SEQ ID NO:7; or

35 c) a nucleic acid sequence that has at least an identity of 80% with the SEQ ID NO:8.

As mentioned above, another embodiment of the present invention  
are plasmid vectors for targeted transformation of filamentous  
40 fungi comprising a PKS-marker. These plasmid vectors are either vectors currently used for targeted transformation of filamentous fungi e.g. such as pAN7 (Punt et al, 1987 Gene 36:117-124) and other vectors that are well known by the skilled artisan or plasmid vectors according to the invention, preferably plasmid vec-  
45 tors according to the invention.

## 24

All of the above-mentioned vectors comprising the PKS marker are hereinbelow termed as "PKS-vectors".

A PKS-vector is also a vector, which comprises a PKS-Marker-expression cassette.

All vectors according to the invention not comprising the PKS marker are hereinbelow termed as "non-PKS-vectors".

10 The present invention furthermore encompasses a method for preparing mutated filamentous fungi, comprising the steps of transferring a non-PKS vector or a PKS-vector into a filamentous fungi; and selecting clones of said filamentous fungi, which contain at least one genetic marker introduced by said plasmid vector.

15

The term filamentous fungi as well as preferred fungi for this purpose are those mentioned and defined above.

In a preferred embodiment, the method for preparing mutated filamentous fungi, comprising the following steps

20

- a) transferring a PKS-vector into a filamentous fungi; and
- b) selecting successfully transformed filamentous fungi by the absence of colour (pigment).

25

As explained above, the absence of colour is based on significant reduction in the amount polyketide synthetase (or in the polyketide synthetase activity or instability of polyketide synthetase mRNA, which is present in a filamentous fungi. The absence of colour can be monitored for example by comparing the transformed fungi with the respective wild-type fungi of the same species.

If a PKS-vector is transferred into a filamentous fungi, the disruption of the PKS gene leads to a loss of colour (pigment) whereby the degree of transformation can be determined easily. Resulting transformants are white in contrast to the colored wild-type. Thus, the selection according to step b) is done by monitoring the absence of melanin in the filamentous fungi. In a preferred embodiment, the absence of pigment is monitored by optical means.

Alternatively, the absence of colour results from the reduction of the polyketide synthetase via antisense techniques. The absence of colour hereby means a "reduction of colour" or, preferably, loss of colour. Absence of colour means a reduction in colour of at least 20%, preferably between 20 and 40%, by prefer-

45

## 25

ence between 40 and 60%, more preferably between 60 and 80%, most preferred between 80% and 100%.

In a more preferred embodiment, the PKS-vector comprises at least  
5 an additional selection marker, preferably the hygromycin resistance gene. In a particular preferred embodiment, the selection of the successfully transformed filamentous fungi comprising a PKS-vector can be carried out by hygromycin resistance of successfully transformed clones and by the absence of pigment of successfully  
10 transformed clones. Most preferably, the PKS-vector is a vector according to the invention additionally comprising a PKS-marker.

In a further embodiment of the invention, the selection of the successfully transformed filamentous fungi comprising a non-PKS-  
15 vector can be carried out by hygromycin resistance of successfully transformed clones.

If a non-PKS-vector is used, the vector is linearized by a restriction enzyme cutting in the nucleic acid sequence region of  
20 element d). Also nucleic acid sequences exceeding 2000 bp can be used what can be disadvantageous as mentioned above. If a PKS-vector is used, the plasmid vector is transferred into a filamentous fungi with the proviso said vector being linearized by a restriction enzyme in PKS-marker nucleic acid sequence. Contrarily to the non-PKS-vectors, the nucleic acid sequence to be  
25 expressed recombinantly can also be smaller than 400bp.

In addition to the aforementioned selection methods set forth in step a) to c), homologous recombination can be confirmed by PCR  
30 based on oligonucleotides preferably derived from the vector sequence flanking the 5 and 3 region of the gene to be inserted. Specific examples of these primers are given in the examples.

The plasmid vector may be transferred into the filamentous fungi  
35 to be transformed by methods familiar to the skilled worker, preferably via protoplast preparation with driselase or driselase and glucanase as lytic enzyme.

The above-mentioned transformation methods can be also realized  
40 in a high throughput screening. Using high throughput screening, many different clones are obtained in parallel so that large numbers of successfully transformed clones of filamentous fungi can be quickly screened.

45 The term filamentous fungi as well as preferred fungi for this purpose are those mentioned and defined above.

## 26

Due to the convenience of the vector, the above-mentioned KO-plasmid preparation, fungi transformation and screening of the mutants can be at least partially automated so that the whole procedure can also be realized in a high throughput screening.

5 Using high throughput system for example for KO-plasmid preparation and DNA amplification by PCR to screen the recombinant mutants, many different clones are obtained in parallel so that large numbers of successfully transformed clones can be quickly screened.

10

Mutagenized filamentous fungi, obtainable according to a method mentioned above, are further encompassed by the present invention.

15 In an alternative embodiment, the method of transforming filamentous fungi based on the use of polyketide synthetase as marker for transformation comprises the following steps:

20 a) providing a filamentous fungi characterized by the absence of colour (pigment), in which the polyketide synthetase gene is modified such that the polyketide synthetase cannot be functionally expressed;

25 b) transforming the filamentous fungi of step a) with a "sense expression cassette" or a vector comprising the aforementioned expression cassette;

c) selecting successfully transformed filamentous fungi by the presence of pigment (colour).

30

The nucleic acid sequence as defined in b) i to v. is herein below termed as PKS encoding sequence.

35 The terms "expression cassette" and "genetic control elements" are explained above.

The "sense-expression cassette" set forth in step b) of the above-mentioned method comprises

40 a) a promotor sequence in functional linkage with a nucleic acid sequence comprising

i. a nucleic acid sequence shown in SEQ ID NO:3, 4 or 5; or

45

## 27

- ii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or
- 5     iii. a functional equivalent of the nucleic acid sequence set forth in i) which is encoded by an amino acid sequence shown in SEQ ID NO:6 that has at least an identity of 40% with the SEQ ID NO:6; or
- 10     iv. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO: 11;
- 15     v. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation;

and optionally

- 20 b) further genetic control sequences in functionally linked with a nucleic acid sequence according to a).

The expression cassette or vector comprises preferably a polyketide synthetase encoding nucleic acid sequence as set forth in b)

25 i., ii. or iii..

Preferred phytopathogenic filamentous fungi are those mentioned above. Preferred non-phytopathogenic filamentous fungi are fungi of group consisting of the genera *Aspergillus* such as *Aspergillus*  
30 *parasiticus*, *Aspergillus nidulans* and *Wangiella* such as *Wangiella dermatidis*.

The modification of the polyketide synthetase encoding sequence of the respective fungi can be done either by introduction of at  
35 least one mutation in the gene encoding a polyketide synthetase or disruption of the gene encoding a polyketide synthetase.

The term "disruption of the PKS marker" means that the PKS marker sequence is disrupted by introducing DNA comprising stop-codons  
40 in the PKS marker sequence e.g. by homologous recombination. The respective methods are well known by the skilled artisan.

The term "mutations" of nucleic acid sequences comprises substitutions, additions, deletions, inversions or insertions of one  
45 or more nucleotide residues, which have to bring about termination of translation of the corresponding amino acid sequence of the target protein by the substitution, insertion or deletion of

one or more amino acids (e.g. a by frame-shift or introduction of stop codon or amendment of nucleic acid sequence). The respective methods are well known by the skilled artisan.

- 5 For example, the mutations are carried out in the flanking regions of exon and intron of a PKS gene. These regions can be determined easily by the skilled artisan. For example, in SEQ ID NO:3 the flanking regions between exon are at bp 1022/1023; bp 1067/1068, bp 1361/1362; bp 1067/1068; bp 1361/1362; bp 1067/1068  
10 ; bp 1361/1362; bp 1416/1417; bp 2399/2400; bp 2447/2448; bp 2675/2676; bp 2738/2739; bp 5744/5745; bp 5792/5793; and/or bp 7205/7206 (Ende 6. exon bp 7205)..

- The term "functional analogues" is defined above describe, in the  
15 presence context nucleic acid sequences which are capable of bringing about the expression, in a filamentous fungi, of a polypeptide with the biological activity of polyketide synthetase and which can be deduced from an amino acid sequence by back translation which has at least an identity of 40%, 41%, 42%, 43%, 44%,  
20 45%, 46%, 47%, 48% or 49% preferred of 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78% or 79% more preferred of 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89% or 90% most preferred of 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or  
25 99% with the SEQ ID NO:6.

- As explained above, the plasmid vector may be transferred into the filamentous fungi to be transformed by methods familiar to the skilled worker, preferably via protoplast preparation with  
30 driselase or driselase and glucanase as lytic enzyme.

- The above-mentioned transformation methods can be also realized in a high throughput screening. Using high throughput screening, many different clones are obtained in parallel so that large numbers of successfully transformed clones of filamentous fungi can be  
35 quickly screened.

- The invention is now illustrated by the examples which follow, but not limited thereto.  
40

#### Examples

- The recombinant methods on which the exemplary embodiments which  
45 follow are based are now described briefly:

## A: General methods

Cloning methods such as, for example, restriction cleavages, DNA isolation, agarose gel electrophoresis, purification of DNA fragments, transfer of nucleic acids to nitrocellulose and nylon membranes, linking of DNA fragments, transformation of *E. coli* cells, bacterial cultures, sequence analysis of recombinant DNA and Southern and Western Blots were carried out as described by Sambrook et al., Cold Spring Harbor Laboratory Press (1989) and Ausubel, F.M. et al., Current Protocols in Molecular Biology, Greene Publishing Assoc. and Wiley-Interscience (1994); ISBN 0-87969-309-6.

The bacterial strains used hereinbelow (*E. coli* DH5 or XL1 blue) were obtained from Life Technologies or Stratagene. The vector were used for cloning. DSM:4527 can be used as *F. Graminearum* wild-type strain 8/1. Restriction maps of the vectors pUCmini-Hyg and PUCmini-Hyg TA are given in Fig 1 and 2.

20

B: Sequence analysis of recombinant DNA (please check, whether this is the method of choice)

Recombinant DNA molecules were sequenced using an ABI laser fluorescence DNA sequencer following the method of Sanger (Sanger et al., Proc. Natl. Acad. Sci. USA, 74, 5463-5467(1977)). Fragments resulting from a polymerase chain reaction were sequenced and verified in order to avoid polymerase errors in constructs to be expressed.

30

## C: Materials used

Unless otherwise specified in the text, all of the chemicals used were obtained in analytical grade quality from Fluka (Neu-Ulm), Merck (Darmstadt), Roth (Karlsruhe), Serva (Heidelberg) and Sigma (Deisenhofen). Solutions were prepared using pure pyrogen-free water, referred to in the following text as H<sub>2</sub>O, from a Milli-Q water system purification unit (Millipore, Eschborn). Restriction enzymes, DNA-modifying enzymes and molecular-biological kits were obtained from AGS (Heidelberg), Amersham (Brunswick), Biometra (Göttingen), Roche (Mannheim), Genomed (Bad Oeynhausen), New England Biolabs (Schwalbach/Taunus), Novagen (Madison, Wisconsin, USA), Perkin-Elmer (Weiterstadt), Pharmacia (Freiburg), Qiagen (Hilden) and Stratagene (Heidelberg). Unless otherwise specified, they were used following the manufacturer's instructions.

## 30

All of the media and buffers used for the genetic engineering experiments were sterilized either by filter sterilization or by heating in the autoclave.

5 In degenerated primer sequences, the following abbreviations are used:

A or T = "W";

G or C = "S";

T or C = "Y";

10 A or C = "M";

A or G = "R";

## Examples

15 Example 1 - Construction of pUCmini-Hyg and PUCmini-Hyg TA vector

A 2536 bp DNA fragment corresponding to the promoter of glycerol-3-phosphate dehydrogenase (GPD1) from *Cochliobolus heterotrophus* associated to the hygromycin B resistance gene from

20 *Escherichia coli* was amplified by PCR with the oligonucleotides

P1 5' atgaagcttgggggtttgagggccaatggaacgaaactagtgtaccacttgacc 3'  
(SEQ ID NO 14); and

25 P2 5'gacagatctggegccattcgccattcag 3' (SEQ ID NO 15)

using pGUS5 as template (Mönke, E. and Schäfer, W., 1993, Mol. Gen. Genet. 241: 73-80). The PCR is done using standard protocols; e.g. as described in Maniatis et al., Mol. Cloning.

30

The resulting DNA fragment was inserted in the plasmid pFDX3809 (WO 01/38504) by the restriction site Hind III and Bgl II introduced by the oligonucleotides P1 and P2. The resulting plasmid pHygB serves as template for a further PCR, wherein the Oligonucleotides

35

ANK 518 5' ggaatcgggtcaatacactac 3' (SEQ ID NO 16)

ANK 519 5' tgtagatctctattcctttgccctcggacgagt 3' (SEQ ID NO 17)

40

are used to shorten the hygromycin B resistance gene specifically. The resulting PCR fragment comprising 575 bp of the 3' end of the hygromycin gene was inserted in the plasmid pHygB via the restriction sites Nde I/ Bgl II generating the plasmid pHygB-NOS.

45

## 31

A Hind III / Ssp I DNA fragment of 2019 bp containing the expression cassette GPD1 promoter, the hygromycine B resistance gene and the nopaline synthase terminator was isolated from pHygB-NOS and inserted in the pUCmini plasmid (= plasmid pFDX3809, see WO 01/38509) previously treated with EcoRI and HindIII restriction enzymes to give the plasmid pUCmini-Hyg; to do so, the EcoRI ends were made compatible with Ssp I by a fill-in treatment using the Klenow fragment of DNA polymerase I. A second version of pUCmini-Hyg, called pUCmini-Hyg-TA, was obtained by the insertion of the following adaptor in the NotI/AscI restriction sites of pUCmini-Hyg:

5' GGCCGCCACGGATATCTTGCCCAAAGAATTCCTGG 3' (SEQ ID NO 18)

15 3' CGGTGCCTATAGAACCGGTTTCTTAAGGACCGCGC 5' (SEQ ID NO 19)

The adaptor contains 2 XcmI restriction sites so that XcmI digest of pUCmini-Hyg-TA creates T-overhangs that permits direct cloning of PCR products made with the classical Taq-polymerases.

20

Example 2 - Construction of the PKS comprising vector "pUCmini-Hyg-PKS"

The nucleic acid sequence encoding PKS was amplified by PCR with 25 degenerated primers

LC1 5'-GAY CCI MGI TTY TTY AAY ATG-3' (SEQ ID NO 20)

LC2c 5'-GTI CCI GTI CCR TGC ATY TC-3' (SEQ ID NO 21)

30

based on the conserved amino acid sequence of the PKS gene sequences from *Aspergillus nidulans*, *Colletotrichum lagenarium*, *Penicillium patulum*, and *Aspergillus parasiticus* (Bingle et al., 1999) using genomic DNA of *Fusarium graminearum* as template.

35 Thermal cycling parameters consisted of an initial denaturation at 94°C for 3 min followed by 34 cycles of 94°C for 1 min (denaturation), 55°C for 1 min (annealing), 72°C for 3 min (extension) and a final extension at 72°C for 10 min according to standard procedures. The resulting PCR product was cloned into the pGEM-T vector (Promega, Mannheim, Germany) to give the plasmid pGEM-T/PKS833 and sequenced. A 633 bp DNA fragment (2236bp to 2870bp of SEQ ID NO:1; corresponding to 2234bp to 2865bp of SEQ ID NO:3; set forth in SEQ ID NO:18) was amplified by PCR using the oligonucleotides

45

ANK593 5' ATAAGAATGCGGCCGCAATGGCCCTCGAAACAGC 3' (SEQ ID NO 22)

## 32

ANK594 5' AAATGGCGCGCCGCGCCAGAAATGACACC 3' (SEQ ID NO 23)

and cloned into the plasmid pUCmini-Hyg using the restriction site NotI and AscI present in the oligonucleotide sequences. The  
5 resulting plasmid pUCmini-Hyg-PKS is used for homologous recombination.

The flanking regions of the PKS DNA fragment were obtained by inverse PCR (Triglia T, Peterson MG, Kemp DJ, Nucleic Acids Res  
10 1988 Aug 25;16(16):8186). Genomic DNA was treated with the restriction enzymes PstI, NcoI, or XhoI respectively. DNA was then self-ligated to get circular DNA molecule. The latter was used as template for the inverse PCR reaction using the primers

15 P1A: 5' TGCCACCTGTAGTCTGCAATCAG 3' (SEQ ID NO 24) and

P2A: 5' TGACTAACCCTGACAACTTCGCTG 3' (SEQ ID NO 25)

deduced from the polyketide synthetase (PKS) DNA fragment of the  
20 plasmid pGEM-T/PKS833 described above.

In a second step, the PCR product was reamplified with the nested primers

25 P1B: 5' CCAGGATCCGACTGCTCAG 3' (SEQ ID NO 26) and

P2B: 5' CTACATCGAGATGCACGGCAC 3' (SEQ ID NO 27)

(deduced from the PKS DNA fragment of the plasmid pGEM-T/PKS833)  
30 , cloned into the pPCR-XL-TOPO vector (Invitrogen) and sequenced to get SEQ ID NO:1.

#### Identification of the genomic DNA Sequence

35 The remaining parts of the flanking regions were obtained by Tail-PCR (Liu YG, Whittier RF; Genomics 1995 Feb 10;25(3):674-81) using 9 arbitrary degenerated primers

40 FJM-tail-AD1	5'-NGT CGA SWG ANA WGA A-3' (SEQ ID NO 28),
FJM-tail-AD2	5'-GTN CGA SWC ANA WGT T-3' (SEQ ID NO 29),
FJM-tail-AD3	5'-WGT GNA GWA NCA NAG A-3' (SEQ ID NO 30),
45 FJM-tail-AD4	5'-NTC GAS TWT SGW GTT-3' (SEQ ID NO 31),
FJM-tail-AD6	5'-TCW GNA GWA NCA SAG A-3' (SEQ ID NO 32),

FJM-tail-AD7 5'-AGW GNA GWA NCA WAG G-3' (SEQ ID NO 33),

FJM-tail-AD8 5'-CAW CGI CNG AIA SGA A-3' (SEQ ID NO 34)  
5 and

FJM-tail-AD9 5'-TCS TIC GNA CIT WGG A-3 (SEQ ID NO 35),

10 coupled to the primer

TailPKS1c 5'-TTG TTA CTG GAG AGG TAA TGA AG-3" (SEQ ID NO 36)

specific for the 5' PKS flanking region deduced from SEQ ID NO:1,  
15 or coupled to the primer

TailPKS2c 5'-TGA GAC AGA TCT CGC GAG CCC TC-3' (SEQ ID NO 37)

specific for the 3' PKS flanking region deduced from SEQ ID NO:1  
20 . After subcloning and subsequent sequencing of the PCR products  
SEQ ID NO:3 was obtained.

#### Identification of the cDNA Sequence of Polyketide Synthetase

25 The PKS cDNA sequence was obtained by RT-PCR with a crude RNA  
preparation from *fusarium graminearum* and various primers de-  
duced from the genomic sequence. This was done according the  
classicals methods (Ausubel, F.M. et al., Current Protocols in  
Molecular Biology, Greene Publishing Assoc. and Wiley-Inter-  
30 science (1994); ISBN 0-87969-309-6). Alignment of cDNA and ge-  
nomic PKS sequences permits to identify precisely the location of  
introns in the genomic sequence.

#### Example 3 Transformation of *F. graminearum*

35

50 ml of CM-medium (Leach et al., 1982, J. Gen. Microbiol. 128:  
1719-1729) were inoculated with approximately  $10^5$  conidia, and in-  
cubated for 2 days at 28°C, 140 rpm. Resulting hyphae were homoge-  
nized in a Warring-Blender; 200 ml CM were inoculated with 10 ml  
40 hyphal suspension, and incubated overnight at 24°C. Mycel were  
trapped on a sterile filter, and washed two times with sterile  
water. 2 g of the hyphae were resuspended in 20 ml Driselase/Glu-  
canase (Interspex Products, San Mateo, USA; 5% / 3% in 700 mM  
NaCl, pH 5.6), and digested 2½ to 3 h at 28°C, 75 rpm. Undigested  
45 hyphal were removed from the protoplast suspension by filtration  
through gauze and Nybold membrane (50 µm pore size). The proto-  
plast suspension were combined with 700 mM NaCl and again passed

## 34

through the gauze and the Nybold membrane. The protoplasts were pelleted by centrifugation (1300 x g) in a swing-out Rotor and washed two times with ice-cold NaCl 700 mM and centrifuge (830 x g). Then the protoplasts were resuspended in STC (0.8 M sorbitol, 50 mM Tris-HCl pH 8.0, 50 mM CaCl<sub>2</sub>) and store on ice until transformation (maximal 1 week).

For transformation, protoplasts were resuspended in 4 parts STC and 1 part SPTC (0.8 M sorbitol, 40% polyethylene glycol 4000, 50 mM Tris-HCl pH 8.0, 50 mM CaCl<sub>2</sub>) at a concentration of 0.5-2 x 10<sup>8</sup>/ml; 30 µg of the pUCmini-Hyg-PKS plasmid DNA linearized with the Eco47III restriction site inside the PKS fragment and 5 µl heparin (5 mg/ml in STC) were added to 100 µl of the protoplast suspension in 10 ml tubes. After mixing, samples were incubated on ice for 30 min. 1 ml SPTC was mixed to the suspension and incubated at room temperature for 20 min. Protoplasts were mixed gently into 200 ml regeneration medium ( 0.1% (w/v) yeast extract, 0.1% (w/v) caseinhydrolysate, 34.2% (w/v) sucrose, 1.6% (w/v) granulated Agar) at 43°C and spread on a 94 mm plates (20 ml per plate). The plates were incubated at 28°C. After 12-24 h, the plates were overlaid with 10 ml per plate water based selective medium (16g/l granulated agar, 100mg/l Hygromycin and further incubated at 28°C until transformants were obtained, which were transferred to fresh CM-Hyg-plates (consisting of CM-media, 100 µg/ml hygromycin and 2% (w/v) Agar. The transformants were isolated by single spore isolation. For generation of conidia, the transformants were cultivated on SNA plates (Nirenberg, 1981, Canadian J. Botany 59: 1599-1609) under UV-light 7-14 days at 18°C. Dilutions of conidia were plated on CM-Hyg plates, and single colonies were transferred from these plates to fresh CM-Hyg plates.

## Example 4 Southernblot analysis

Genomic DNA was isolated from frozen hyphal material using the Puregene Genomic DNA Isolation Kit (Gentra Systems, Minneapolis USA) and digested for 6 h with NruI restriction enzyme. The genomic DNA was separated by electrophoresis on a 1% (w/v) agarose gel and blotted onto a nylon membrane (Hybond NX; Amersham Pharmacia Biotech, Buckinghamshire England). A digoxigenin labeled probe was generated by PCR based on specific primers PKS forward 5'-GCG CTT GAG ATG GCT AGT ATC G-3' and and PKS reverse 5'-GTG CCG TGC ATC TCG ATG TAG-3' using pGEM-T/PKS833 as template and digoxigenin labeled dUTPs by PCR reaction according to the recommendation of the manufacturer(Roche Diagnostics GmbH, Mannheim). PCR conditions were 94°C for 3 min (initial denaturation) followed by 30 cycles of 94°C for 30 sec (denaturation), 55°C for 45 sec (annealing),

## 35

72°C for 1 min (extension) and a final extension at 72°C for 10 min. The non-radioactive hybridization and the detection were done under highly stringent conditions as described in Roche Molecular Biochemicals DIG Application Manual for Filter Hybridization (Roche Diagnostics GmbH, Mannheim).

To confirm the insertion of the vector construct into the PKS locus in comparison with the wild type gene, primers

10 EF-PKS 5' atgtctccaaaggaagctgagc 3' (SEQ ID NO 38); and

ER-PKS 5' tcgagtgatggatactgcttcg 3' (SEQ ID NO 39)

are constructed based on the PKS DNA sequence from the plasmid 15 pGEM-T/PKS833; four universal primers are constructed, wherein

Lac 92 5' cggctacactagaaggacagtatttggtta 3' (SEQ ID NO 40)

20 Lac 93 5' gtcaggcaactatggatgaacgaaatagac 3' (SEQ ID NO 41)

Lac 94 5' acccatctcataaataacgtcatgc 3' (SEQ ID NO 42); and

Lac 95 5' caactctatcagagcttggttga 3' (SEQ ID NO 43)

25 permit amplification of a 412 bp DNA fragment of the hygromycin cassette.

PCR reactions were conducted in classical conditions: 94°C for 3 min (initial denaturation) followed by 30 cycles of 94°C 60 sec 30 (denaturation), 55°C for 90 sec (annealing), 72°C for 90 sec (extension) and a final extension at 72°C for 10 min.

6 recombinant clones resistant to Hygromycine were analyzed by PCR using the primer set Lac 94 / Lac 95 specific for the hygromycin resistance gene. All the mutants were found to present the 35 expected DNA fragment of 412 bp indicating the integration of the plasmid pUCmini-Hyg-PKS in the genome.

A 712 bp corresponding to the PKS gene could be amplified with the primer set EF-PKS/ER-PKS mentioned above using genomic DNA 40 from a wild type strain; on the contrary no PCR fragment were amplified with genomic DNA from the recombinant clones indicating that the PKS gene is disrupted by the insertion of pUCmini-Hyg-PKS. This was confirmed by PCR amplification EF-PKS combined with Lac 93 (hybridizing to the plasmid backbone near Not I restriction 45 site) and ER-PKS combined with Lac 92 (hybridizing to the plasmid backbone near Asc I restriction site). In both cases, DNA fragments of about 600 bp were amplified for the recombinant

## 36

clones but not for the wild type strain (WT). All together the PCR analysis using the different primer sets proves that the plasmid pUCmini-Hyg-PKS was targeted specifically in the PKS locus by homologous recombination. This process permits to disrupt  
5 the PKS gene since the recombinant mutants were found to lack the typical pigmentation (purple) of the wildtype strain.

Example 5 functional expression of Green Fluorescent Protein (GFP) in *Fusarium graminearum*

10

A) Plasmid construction

In a first step, a 67bp DNA fragment encoding the peptide leader of the first 23 amino acids from N-terminus of the yeast ARH1  
15 (SwissProt; P48360) was amplified by PCR using the primers

Lac 80 5' cccgaattcatgagctttgttcaaataagg 3' (SEQ ID NO 44) and

Lac 81 5' ttattctagattttccatgggaatggatacagtcttacg 3' (SEQ ID NO  
20 45)

In a second step, a 734 bp DNA fragment encoding the Green Fluorescent Protein (GFP) was amplified by PCR using the plasmid pEGFP-N2 (Genbank; U57608) and the primers

25

Lac 84 5' cgccaccatggtgagcaagggcgaggagctgtt 3' (SEQ ID NO 46) and

Lac 85 5' tatgatctagagtcgcggccgctttacttgtacagctcg 3' (SEQ ID NO  
30 47).

The PCR products were assembled in frame with the Nco I restriction sites present in the oligonucleotides Lac 81 and Lac 84 and cloned in the expression plasmid pYes2 (Invitrogen) using the restriction sites EcoRI and Xba I present in the oligonucleotides  
35 Lac 80 and Lac 85, respectively. In the resulting plasmid pLAC7, the recombinant gene encoding GFP is under the control of the galactose (Gal 1) promoter and cytochrome C1 terminator.

40 A 2892 bp DNA fragment containing the GFP expression cassette was isolated from pLac7 using the restriction sites Nae I and Bsa I and cloned in the plasmid pUCmini-Hyg-PKS (see example 2). To do so, pUCmini-Hyg-PKS was firstly cut by Asc I and filled in according to classical methods then treated with Bsa I. The re-  
45 sulting plasmid pUCmini-Hyg-PKS-GFP contains all genetic elements

permitting the production of recombinant GFP in *Fusarium graminearum*.

B) Transformation of *Fusarium graminearum* with pUCmini-Hyg-PKS-5 GFP and analysis of transformants

The transformation was done as described in example 3, wherein pUCmini-Hyg-PKS-GFP was linearized with EcoR47III. The correct integration of the plasmid in the PKS locus was observed after 10 single conidiation by the absence of pigmentation of the recombinant mutants.

In addition, the integration was confirmed by PCR as described in example 4 using the following primer combinations EF-PKS (see ex- 15 ample 4; SEQ ID NO:38) and ER-PKS (see example 4; SEQ ID NO:39), whereby no amplification were observed since the gene PKS is disrupted whereas wild type strain or unspecific mutants were presenting a 714 bp DNA fragment corresponding to the expected PKS DNA fragment.

20

Using the primer combination EF-PKS (see example 4; SEQ ID NO:38) and

Lac 211 5' gcttctaataccgtactagtgatca 3' (SEQ ID NO 48)

25

the amplification of a 835 bp DNA corresponding to the 5' end plasmid integration in the PKS locus of the mutants was observed. No DNA fragment was amplified for the wildtype strain due the absence of the DNA sequence complementary to Lac211.

30

The primer combination

ANK 458 5' ctttgatcttttctacgggggtctga 3' (SEQ ID NO 49) and

35 ER-PKS (see example 4; SEQ ID NO:39) led to the amplification of a 718 bp DNA corresponding to the 3' end plasmid integration in the PKS locus of the mutants. No DNA fragment was amplified for the wildtype strain due the absence of the DNA sequence complementary to ANK 458.

40

C) Detection of the production of GFP

The recombinant mutants were grown for a few days in CM-Hyg medium as described in example 3 except for glucose which was re- 45 placed by galactose as a carbon source. The fluorescence of GFP was detected using the polarstar spectrophotometer (Firma BMG; Ex: 385nm and Em: 520nm). In these conditions fluorescence was

## 38

observed for the strains which showed integration of the plasmid whereas no fluorescence was observed for the wildtype strains.

Brief description of the figures

5

Figure 1: Map of pUCmini-Hyg

Figure 1: Map of PUCmini-Hyg TA

10

15

20

25

30

35

40

45

## Claims

1. A plasmid vector for targeted transformation of filamentous  
5 fungi comprising
  - a) an origin of replication for a host organism which is not  
originating from the filamentous fungi to be transformed;
  - 10 b) a selection marker for a host organism not originating  
from the filamentous fungi;
  - c) a promotor facilitating recombinant expression in fungi  
15 that is functionally linked to the coding region of the  
hygromycin resistance gene which is functionally linked  
to a terminator which facilitates transcription termina-  
tion in filamentous fungi;
- wherein the overall size of the elements a), b) and c) does  
20 not exceed 4500 bp; and
  - d) a nucleic acid sequence, which is homologous to nucleic  
acid sequences of the filamentous fungi to be transformed  
and makes homologous recombination in the filamentous  
25 fungi to be transformed possible.
2. A plasmid vector as claimed in claim 1, wherein the an origin  
of replication a) originates from bacteria.
- 30 3. A plasmid vector as claimed in claim 1 to 2, wherein the  
selection marker b) imparts a resistance to antibiotics.
4. A plasmid vector according to claim 1 to 3, wherein the pro-  
motor of element c) is selected from the group consisting of  
35 the GPD-1-, PX6-, TEF-, CUP1-, PGK-, GAP1-, TPI, PHO5-, AOX1,  
GAL10/CYC-1, CYC1, OliC-, ADH-, TDH-, Kex2-, MFa- and the  
NMT-promotor.
5. A plasmid vector according to claim 1 to 4, wherein the ter-  
minator of element c) is selected from the group consisting  
40 of the AOX1-, nos-, PGK-, TrpC- and the CYC1-terminator.
6. A plasmid vector according to claim 1 to 5, wherein the pro-  
motor of element c) is the GPD-1-promotor and the terminator  
45 of element c) is the nos-terminator.

## 40

7. A plasmid vector according to claims 1 to 6, wherein the nucleic acid sequence d) is functionally linked to a promotor facilitating recombinant expression in filamentous fungi.
- 5 8. A plasmid vector according to claims 1 to 7, wherein the nucleic acid sequence d) is functionally linked to a transcription terminator facilitating recombinant expression in filamentous fungi.
- 10 9. A selection marker comprising a nucleic acid sequence encoding a polyketide synthetase fragment, wherein said nucleic acid sequence comprises
- 15 i. a nucleic acid sequence shown in SEQ ID NO:1 or SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4 or SEQ ID NO:5; or
- ii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or
- 20 iii. a functional equivalent of the nucleic acid sequence set forth in i) which is encoded by an amino acid sequence that has at least an identity of 50% with the SEQ ID NO:6; or
- 25 iv. parts of the nucleic acid sequence as defined in i., ii. or iii. consisting of at least 300bp; or
- 30 v. parts of the nucleic acid sequence as defined in i., ii. or iii. consisting of at least 300bp comprising
- a) a nucleic acid sequence shown in SEQ ID NO:7 ; or
- 35 b) a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:8 by back translation; or
- 40 c) a functional equivalent of a nucleic acid sequence set forth in a), which is encoded by amino acid sequence that has at least an identity of 85% with the SEQ ID NO:8.

10. Use of a nucleic acid sequence comprising

## 41

a) a nucleic acid sequence encoding a polyketide synthetase;  
or

5 b) parts of the nucleic acid sequence as defined in i. consisting of at least 300bp.

as marker for targeted transformation in filamentous fungi.

11. Use of a nucleic acid sequence according to claim 10 said nucleic acid sequence comprising

i. a nucleic acid sequence according to claim 9; or

15 ii. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO: 11; or

20 iii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation; or

25 iv. a functional equivalent of the nucleic acid sequence set forth in i), which is encoded by an amino acid sequence that has at least an identity of 40% with the SEQ ID NO:6; or

v. parts of the nucleic acid sequence as defined in ii., iii. or iv. consisting of at least 300bp; or

30 vi. parts of the nucleic acid sequence as defined in ii., iii or iv. consisting of at least 300bp comprising a nucleic acid sequence, which is encoded by an amino acid sequence that has at least an identity of 68% with the SEQ ID NO:8.

35

12. A plasmid vector for targeted transformation of filamentous fungi additionally comprising a selection marker comprising a nucleic acid sequence encoding a polyketide synthetase fragment said nucleic acid sequence comprising

40

i. a nucleic acid sequence according to claim 9; or

45 ii. a functional equivalent of the nucleic acid sequence set forth in i) which is encoded by an amino acid sequence that has at least an identity of 40% with the SEQ ID NO:6.

## 42

- iii. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO: 11;
- 5 iv. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation; or
- 10 v. parts of the nucleic acid sequence as defined in i., iii. or iv. consisting of at least 300bp; or
- 15 vi. parts of the nucleic acid sequence as defined in i., ii. or iii. or iv. consisting of at least 300bp, which are encoded by an amino acid sequence that has at least an identity of 68% with SEQ ID NO:8.
13. A plasmid vector for targeted transformation of filamentous fungi as claimed in claims 1 to 8, additionally comprising a selection marker comprising a nucleic acid sequence encoding a polyketide synthetase fragment said nucleic acid sequence comprising
- 20 i. a nucleic acid sequence according to claim 9; or
- 25 ii. a functional equivalent of the nucleic acid sequence set forth in i), which is encoded by an amino acid sequence that has at least an identity of 40% with the SEQ ID NO:6; or
- 30 iii. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO: 11;
- 35 iv. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation; or
- 40 v. parts of the nucleic acid sequence as defined in i., iii. or iv. consisting of at least 300bp; or
- 45 vi. parts of the nucleic acid sequence as defined in i., ii. or iii. or iv. consisting of at least 300bp comprising a nucleic acid sequence, which is encoded by a functional equivalent of an amino acid sequence that has at least an identity of 68% with the SEQ ID NO:8.

14. An expression cassette comprising

5 a) a promotor sequence in functional linkage with a nucleic acid sequence according to claim 9 in antisense orientation; and optionally

b) further genetic control sequences in functionally linked with a nucleic acid sequence according to a).

10 15. A plasmid vector for targeted transformation of filamentous fungi additionally comprising an expression cassette according to claim 14.

15 16. A plasmid vector for targeted transformation of filamentous fungi as claimed in claims 1 to 8, additionally comprising an expression cassette according to claim 14.

17. A method for transforming filamentous fungi, comprising the following steps

20

a) transferring a plasmid vector according to claim 12, 13, 15 or 16 into a filamentous fungi;

25

b) selecting successfully transformed filamentous fungi by the absence of colour.

18. An expression cassette comprising

30

a) a promotor sequence in functional linkage with a nucleic acid sequence comprising

i. a nucleic acid sequence shown in SEQ ID NO:3, 4 or 5;  
or

35

ii. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:6 by back translation; or

40

iii. a functional equivalent of the nucleic acid sequence set forth in i) which is encoded by an amino acid sequence shown in SEQ ID NO:6 that has at least an identity of 40% with the SEQ ID NO:6; or

45

iv. a nucleic acid sequence shown in SEQ ID NO:9 or SEQ ID NO: 11;

## 44

- v. a nucleic acid sequence which, owing to the degeneracy of the genetic code, can be deduced from the amino acid sequence shown in SEQ ID NO:10, SEQ ID NO:12 or SEQ ID NO:13 by back translation;

5

and optionally

- b) further genetic control sequences in functionally linked with a nucleic acid sequence according to a).

10

19. A method for transformation of filamentous fungi, comprising the following steps

15

- a) providing a filamentous fungi, in which the polyketide synthetase gene is modified such that the polyketide synthetase cannot be functionally expressed;

20

- b) transforming the filamentous fungi of step a) with an expression cassette according to claim 18 or vector comprising the aforementioned expression cassette;

- c) selecting successfully transformed filamentous fungi by the presence of colour.

- 25 20. A method as claimed in claim 17 or 19, wherein the plasmid vector comprises at least an additional selection marker.

21. A method as claimed in claims 17, 19 or 20, wherein the selection is confirmed by PCR.

30

22. A method as claimed in claims 17, 19, 20 or 21, wherein the filamentous fungi are successfully transformed and identified in a high-throughput screening.

35

40

45

## Plasmid vectors for transformation of filamentous fungi

## Abstract

5

The invention relates to novel plasmid vectors for transformation of filamentous fungi and to a method of modifying the genome of filamentous fungi based on these vectors. The invention furthermore relates to the modification of a specific gene via the process of homologous recombination, to recombinant expression of foreign genes in filamentous fungi and to new selection markers for detecting successful transfer of the target gene in filamentous fungi.

15

20

25

30

35

40

45

## 1

## SEQUENCE LISTING

&lt;110&gt; BASF Aktiengesellschaft

&lt;120&gt; Plasmid vectors for transformation of filamentous fungi

&lt;130&gt; Polyketide Synthetase

&lt;140&gt;

&lt;141&gt;

&lt;160&gt; 49

&lt;170&gt; PatentIn Ver. 2.1

&lt;210&gt; 1

&lt;211&gt; 5888

&lt;212&gt; DNA

&lt;213&gt; Fusarium graminearum

&lt;400&gt; 1

```
atcgagtttc gtgttgcggtg tcatgggtgta gogaacaata ccatactcct ccatcaaagg 60
gctgacgagc ttggcatgga cctttgtcat atgggtgcctg tactctttct cgctaagaga 120
agg-ttttttg tagggcgtga tggtaggca aacaagggtg tcagggtgac cgggtgacagt 180
tgccatcatga gaggtttctt gggatcact atgaggcgac atgatgcagg ttggagtttg 240
ttatcgaagt gctgggtagg atttggtga cagagttgca taaactggag ctctggatac 300
gcaaggggtc ataagagtat ggatacgccg tcgattagaa gctgtcagat ccagtgcac 360
aatggccoga atgacgagag attgctgctc ggcactgaaa gtctgcagcg tgcttgtttc 420
ggagtatatc ctacagctgc accgaatgac gaacaagccc gactatccat gtccgacgag 480
cccggaccga cgtgatccga atgctgtcta acacttacca agacaatcat ggggtatctc 540
tgaattatga ttgtgttgct caagcagcta atoccaaactc ttaatcacgt caagtgccat 600
gcttacctgt attatcatct gtcacaaaca taccaaactcc tatatatagg ggatttgtgt 660
tcattacctc tccagtaaca actacttggg gttaaacca catttcaact tactctttta 720
gttcattcaa catgacccca tcaatgatgg aggtattcgt ttttggggac caaagcacac 780
gctttgcccc tccactgaaa gacctactcc tcaaaggcaa cagtccttac ttgacacatt 840
ttgttaaaca agttcacgca cttcttagaa gggagatata atccttgccg gcagttcaac 900
agaagctttt cccaaacttt gccgacattc aggaactcgt ctccaagtca gattggggca 960
gtggttaacc ttgctttgaca agcgttttag catgctttta ccactcttgc agtttcattc 1020
agttagtaac ataaccaagt acccatagat ctctcctgac catccagcct ttacagatgg 1080
acaaggctgt accttctctt cggagaacag tcgcattatt ggactttgag ttggttcaact 1140
cgctgctact gctgtcagtt gctccacatc actgagttaa ttggtatcag ctggtgtaga 1200
tgctgttcgt gtggcattgc acgtcggact acgggtatgg cgaactacct ccccttttga 1260
tgtaccagac aggcctctcg ccaactgggt cataattgtg cccgaggcag tactaccaag 1320
agaatctgag caagaccgac ttgactcatt catcattgaa atggttaagca attctttgtc 1380
cttggagctc tcataaaaaca tattcaactc tcttaggga cttgctcgat catcagttcc 1440
ttacatcagc tcgggtcgac atcacaacat gaccatcagt ggtccaccat ccgtcctcga 1500
aaagttcatt cacagtatat caacatcacc gaaagattct cttccagtgc cgatctatgc 1560
tcogtaccac gccagccatc ttacagcat ggatgatgta gacgaggtcc ttagcctgtc 1620
tgcaccttct ttgcatcag agtccatcat tccactcatt tcaagctcct cgggtgacga 1680
gttacagcca ctcaagtatg cagatctact ccgctgctgt gttagtata tgctcatata 1740
gccactggat cttaccaagg tctcacaagc agtggccag cttctcgagg ttagctcatc 1800
tacacgtgcc ataataaagc ctatagcaac cagcgtctcc aacagtctag tgtctgtttc 1860
ggagccgagc ctagcagaac gatgcgccgt ggacaacagc atggggccca aagcctcgac 1920
```

2

cagccactca tcagcagaga cacaaccga gtcacaaagc aagaactcca aaattgcat 1980  
 tgttgctatg tctggcgct ttcagacgc agctgacttg agtgaattct gggatcttct 2040  
 ctacgaaggc cgcgatgttc atcgacaaat tcccaggac cgattcaacg cagagctcca 2100  
 ttaacagcct actgggagac gtaagaacac cagcaaggct atgaatggct gcttcatcaa 2160  
 ggaaccagga ctgttcgacg ctagggttctt caacatgtct ccaaagggaag ctgagcagtc 2220  
 ggatcctggc cagcgaatgg ccctcgaaac agcttacgag gcgcttgaga tggctagtat 2280  
 cgtaccagac agaacacctt cgacacagag agatcgtgtt ggtgtgttct acggcatgac 2340  
 tagcgatgat tggagagagg tcaacagtgg gcagaatgtc gacacttatt ttattcctgg 2400  
 taagtgtttt gttcgcgttt cagacaata ctgattgcag actacagggtg gcaacagagc 2460  
 gttcactcct ggtcgactca actacttctt caagtccagt gggcctagcg ctagtgttga 2520  
 tacggcttgc tctccagtc tegtggcctt gcacttggct tgaattccc tctggagaaa 2580  
 tgattgcat acagcattg cgggcggaac caatgtcatg actaaccttg acaacttcgc 2640  
 tggtttggac cgaggccact tctatctca agaaccggtt agtctctcac agatcttggg 2700  
 agcttcttct tttttagtgt ttttactgac cttgttccag gcaactgcaa caccttgac 2760  
 gatggagcag acggatactg tcgagctgat ggcgtcggaa ccatcatcct caagcggctt 2820  
 gaggacgccc aagctgacaa tgacctatt ctgggtgtca ttctgggccc ttacacaaac 2880  
 cactcagccc aagcagtcac catcactcga ccacatgccg gagctcaaga gtacatcttc 2940  
 tccaaactcc tccgtgagtc gggcaccgat cctacaacg ttagctacat cgagatgcac 3000  
 ggcacaggca ctcaagccgg cgacgcaacc gagatgacat ccgtcctcaa gacgtttgct 3060  
 cctaccagcg gcttcggcgg tcgattgcct caccaaaacc ttcacttggg ttcagtcaag 3120  
 gccaatgtcg ggcacgggtga atccgcctct ggtatcattg ctctgatcaa gacgtgctt 3180  
 atgatggaga agaacatgat cccgccgcat tgtggtatca agacaaagat caatcaccat 3240  
 tttcctacgg atctcactca gcgcaatgtc catatcgcca aagtccgac atcttggaca 3300  
 agatcgggtc aagccaatcc acgcattgct ttcgtcaata acttctctgc cgttgggtgt 3360  
 aactctgctg tctactgca agatgtcct cagccatcgg tagttccgga tgttacagac 3420  
 cctcgacat cccatgttgt cactatgtcc gctcgatcag cagattcctt caggaagaac 3480  
 ctgcacaat tcaaggagct tgtagaaggc caagggtgact cggaggtcgg ctctctgagc 3540  
 aagctgtcct acacaaccac cgccaggcgc atgcatcatc aattccgagc ttcggtcaca 3600  
 gcacagactc gtgaacagct gctgaagggc cttgatcccg ccattgaacg ccaggatgtg 3660  
 aagaggatcc ccgcgcgcgc gccctctgtc ggctttgtgt ttagcggcca aggcgccccaa 3720  
 tacggtggta tgggcaagga gtactttaca tctttcacag ccttccgctc tgagatcatg 3780  
 tcttacgaca gtatgcccga agcccaaggc tcccgtcaa tctccact gatccgagga 3840  
 gaggtggaag ctgactcgtt gagtctgtt gagatccagc tgggtctcac ttgcctgcag 3900  
 atggcactgg ccaagctatg gaagtcattc ggtgttgagc caggtcttgt tctcgacac 3960  
 agcttaggoc actatgctgc ttacacgtc gctgggtgtt tgtccgcaa tgataccatt 4020  
 tactcactg gcacagagc gcagctgctc gtggataagt gccaggcagg aaccactca 4080  
 atgctggcag tgagggcac cttactacag atccaacagt tctcgatgc caacattcac 4140  
 gaggttgcac gtgtcaatgg ttcacgagaa gtcgtcatca gtggacgcgt tgccgacatt 4200  
 gaccagctgg ttggcctttt gtcggctgac aacatcaagg cgaccgcgt caaggtgcca 4260  
 tttgcttcc actcagcga ggttgacccc attctctccg acttggatac agcggcgtcg 4320  
 cgcgtcacct tccactcctt ccagattcct gttctttgtg ccttgacag ctctgtcatc 4380  
 agcctggaa accatggtgt cattggctcc cttcatctac agcgacattg tctgagaca 4440  
 gtcaactttg aggtgtctt acatgctgcg gagcacgaga agatcatcaa caagacatca 4500  
 actctatgga tcgagattgg tccccatgtt gtctgtctca ccttctcaa gtccagcctt 4560  
 ggtccaagea cactgctat cgcacgctt cgcgaaatg acgattgctg gaaggtgttg 4620  
 gctgatgggt tgagcagct ctacagcagc gggttgacaa ttgactagaa cgagtatcat 4680  
 cgcgacttca aggcctctca ccaggtaact cgtctgcctt gttacagctg ggagcacaag 4740  
 aattactgga tacagtacaa gtacgattgg tcttggcta aaggtgatcc tccaattgac 4800  
 cctaacagct cggttgaagc agtctcagct ttatcaacac cctcggctca gaagattcta 4860

3

caggagactt cccttgatca ggtattgact atcgttgctg agacagatct cgcgagccct 4920  
 ctattgtcag aggttgccca aggtcatcgg gtcaacgggtg tcaaagtctg cacatcttcc 4980  
 gtgtacgctg atgttggctt gacgctgggt aagtacattt tggacaacta ccgcaccgac 5040  
 ttagagggtt atgcggctga tgttcacgggt attgagggtcc acaagccact tcttctcaaa 5100  
 gaagacatga acggaacgce ccaggctaca ccgttccgca tcgaagtgcg atacccaate 5160  
 cagagcacca cggcgctgat gagcatctcc accactggcc ccaacgggtca gcacatcaag 5220  
 catgctaact gcgaacttcg actcgagcat ccgtcgcaat gggaagcgga gtgggatcgc 5280  
 caagcctacc tcatcaatcg cagcgtcaac tgccttctgc agcgatcagc acaagggttg 5340  
 gacagcatgt tggcaaccgg aatggtctac aagggtcttct cctccctcgt cgactatgcc 5400  
 gatggctaca aggggtctgca ggagggtgtc ttgcacagcc aagagctcga gggcacagca 5460  
 aaagtgcgct tccaaactcc ctcgggagggt ttcgctctgca atcccatgtg gattgacagc 5520  
 tgtggctcaga cgaccggctt catgatgaac tgtcatcaga ctacgcccga tgactacgtc 5580  
 tacgtcaate atggctggaa gtcgatgaga ttggccaagg cgttcctgtga agatgggtacc 5640  
 tatcggaactt atatccggat gagggccatt gatagcacca agttcgctgg tgacttgtac 5700  
 attcttgatg aggatgacac tgtgggttgggt gtttatggag acataacaat gagtaacacg 5760  
 aatctttgtc aacaacaatg ctaacgtaaa gaacagttcc aagggttgcc gcgacgagtt 5820  
 ctcaacacag tcttgccatc tgccaacgcg gttccagttg atgctcccat gggtcaagac 5880  
 gacctcca 5888

&lt;210&gt; 2

&lt;211&gt; 5157

&lt;212&gt; DNA

<213> *Fusarium graminearum*

&lt;400&gt; 2

atgaccccat caatgatgga ggtattcggt tttggggacc aaagcacacg ctttgcccct 60  
 ccactgaaag acctactcct caaaggcaac agtccttact tgacacattt tgttaaacia 120  
 gtacacgcac ttcttagaag ggagatatca tccttgccgg cagttcaaca gaagcttttc 180  
 cczaactttg ccgacattca ggaactcgtc tccaagtcag attggggcag tggtaacct 240  
 gctttgacaa gcgcttttagc atgcttttac catctttgca gtttcattca gtgagtaaca 300  
 taaccaagta cccatagatc tctcctgacc atccagcctt ttacgatgga caaggctcgt 360  
 cct.ttccttc ggagaacagt cgcattattg gactttgcgt tggttcactc gctgctactg 420  
 ctgtcagttg ctccacatca ctgagtgaat tggatcage tgggtgtagat gctgttcgtg 480  
 tggcattgca cgtcggacta cgggtatggc gaactacctc ccttttcgat gtaccagaca 540  
 ggccctccgc cacttggttc ataattgtgc ccgaggcagt actaccaaga gaatctgcgc 600  
 aagaccgact tgactcattc atcattgaaa tggtaagcaa ttctttgtcc ttggagctct 660  
 cataaaacat attcactcct ctttagggac ttgctcgatc atcagttcct tacatcagct 720  
 cggtcgcaca tcacaacatg accatcagtg gtccaccatc cgtcctcgaa aagttcattc 780  
 acagtataac aacatcaccg aaagattctc ttccagtgcc gatctatgct ccgtaccacg 840  
 ccagccatct ttacagcatg gatgatgtag acgaggctct tagcctgtct gcaccttctt 900  
 ttgcatcaga gtccatcatt ccactcattt caagctcctc gggtagcgag ttacagccac 960  
 tcaagtatgc agatctactc cgctgctgtg ttagtgatat gtcatacag ccactggatc 1020  
 ttaccaaggt ctcaacaaga gtggcccagc ttctcgaggt tagctcatct acacgtgcca 1080  
 taataaagcc tatagcaacc agcgtctcca acagtctagt gtctgttttg gagccgacgc 1140  
 taycagaacg atgcgccgtg gacaacagca tggggcccaa agcctcgacc agccactcat 1200  
 cagcagagac acaaacccgag tcatcaagca agaactccaa aattggcatt gttgctatgt 1260  
 ctgggtcgctt tccagacgca gctgacttga gtgaattctg ggatcttctc tacgaaggte 1320  
 gcgatgttca tcgacaaatt cccgaggacc gattcaacgc agagctccat tacgacgcta 1380  
 ctgggacgag taagaacacc agcaaggctc tgaatggctg cttcatcaag gaaccaggac 1440  
 tgttcgacgc taggttcttc aacatgtctc caaaggaagc tgagcagtcg gatcctggcc 1500  
 agcgaatggc cctcgaaaca gcttacgagg cgcttgagat ggctagtatc gtaccagaca 1560

gaacaccttc gacacagaga gatcgtgttg gtgtgttcta cggcatgact agcgatgatt 1620  
 ggagagaggc caacagtggg cagaatgtcg acacttattt tattcctggg aagttgtttg 1680  
 ttccgctttc acgacaatac tgattgcaga ctacaggtgg caacagagcg ttcactcctg 1740  
 gtcgactcaa ctacttcttc aagttcagtg ggcctagcgc tagtggtgat acggccttgc 1800  
 cctcagtcct cgttggcttg cacttggctt gtaattccct ctggagaaat gattgcgata 1860  
 cagctattgc gggcggaacc aatgtcatga ctaaccctga caacttcgct ggtttggacc 1920  
 gaggccactt cctatctcaa gaaccggtaa gtctctcaca gatcttggaa gcttccttct 1980  
 ttttagtggt ttacttgacc ttgttccagg caactgcaac acctttgacg atggagcaga 2040  
 cggatactgt cgagctgatg gcgtcggaa catcatcttc aagcggcttg aggacgcga 2100  
 agctgacaat gaccctattc tcgggtgtcat tctgggctgt tacacaaacc actcagccga 2160  
 agcagtatcc atcactcgac cacatgccgg agctcaagag tacatcttct ccaaactcct 2220  
 ccgtgagtcg ggcaccgatc cctacaacgt tagctacatc gagatgcacg gcacaggcac 2280  
 tcaagccggc gacgcaaccg agatgacatc cgtcctcaag acgtttgctc ctaccagcgg 2340  
 cttcggcggc cgattgcctc accaaaacct tcacttgggt tcagtcaagg ccaatgtcgg 2400  
 gcacggtgaa tccgcatctg gtatcattgc tctgatcaag acgctgctta tgatggagaa 2460  
 gaacatgata ccgcgcatt gtggtatcaa gacaaagatc aatcaccatt ttcctacgga 2520  
 tctcactcag cgcaatgtcc atatcgccaa agttccgaca tcttggacaa gatcgggtca 2580  
 agccaatcca cgcattgctt tcgtcaataa cttctctgcc gctgggtggt actctgctgt 2640  
 cctactgcaa gatgtcctc agccatcggt agtttcggat gttacagacc ctgcacatc 2700  
 ccatgttgtc actatgtccg ctcgatcagc agattcctc aggaagaacc tcgccaatct 2760  
 caaggagctt gtagaaggcc aaggtgactc ggaggtcggc ttcttgagca agctgtccta 2820  
 cacaaccacc gccaggcgca tgcacatca attccgagct tcggtcacag cacagactcg 2880  
 tgaacagctg ctgaagggcc ttgattccgc cattgaacgc caggatgtga agaggatccc 2940  
 cgcgcgcgcg cctctgtcg gcttctgtgt tagcggccaa ggcgccaat accgtggtat 3000  
 gggcaaggag tactttacat ctttcacagc ctccgctct gagatcatgt cttacgacag 3060  
 tatcgcccaa gcccaaggct tcccgtaaat cctccactg atccgaggag aggtggaagc 3120  
 tgaactggtg agtctgttg agatccagct gggctcact tgcctgcaga tggcactggc 3180  
 caagctatgg aagtcattcg gtgttgagcc aggtttgtt ctcgacaca gcttaggcca 3240  
 ctatgctgct ttacacgtcg ctggtgttct gtccgcaat gataccattt acctactgg 3300  
 catcagagcg cagctgctcg tggataagtg ccaggcagga acccactcaa tgctggcagt 3360  
 gagggcattc ttactacaga tccaacagtt cctcgatgcc aacattcacg aggttgcatg 3420  
 tgtcaatggt tcacgagaag tcgtcatcag tggacgcgt gccgacattg accagctggt 3480  
 tggccttttg tcggctgaca acatcaaggc gaccgcgtc aaggtgccat ttgccttcca 3540  
 ctacagcgag gttgaccca ttctctccga cttggataca ggcgcgtcgc gcgtcacctt 3600  
 ccactccctc cagattcctg ttctttgtgc ccttgacagc tctgtcatca gccctggaaa 3660  
 ccatgggtgc attggtcccc ttcatctaca gcgacattgt cgtgagacag tcaactttga 3720  
 ggggtgctcta catgctgcgg agcacgagaa gatcatcaac aagacatcaa ctctatggat 3780  
 cgagattggt ccccatggtg tctgctctac cttcctcaag tccagccttg gtccaagcac 3840  
 acctgctatc gcacgcctc gccgaaatga cgattgctgg aaggtgttgg ctgatggttt 3900  
 gagcagtcct tacagcagcg ggttgacaat tgactagaac gagtatcatc gcgacttcaa 3960  
 ggcctctcac caggtaactc gtctgccttg ttacagctgg gagcacaaga attactggat 4020  
 acagtacaag tacgattggt ccttggctaa aggtgatcct ccaattgccc ctaacagctc 4080  
 ggttgaagca gtctcagctt tatcaacacc ctcggtccag aagattctac aggagacttc 4140  
 ccttgatcag gtattgacta tcgttgctga gacagatctc gcgagccctc tattgtcaga 4200  
 ggttgcccaa ggtcatcggg tcaacggtgt caaagtctgc acatcttccg tgtacgtga 4260  
 tgttggcttg acgctgggta agtacatttt ggacaactac cgcaccgact tagagggtta 4320  
 tgccgtcgat gttcacggtt ttgaggtcca caagccactt cttctcaaag aagacatgaa 4380  
 cggaacgcc caggctacac cgttcgcat cgaagtgcga racccaatcc agagcaccac 4440  
 ggcgctgatg agcatctcca ccaactggccc caacggctcag cacatcaagc atgctaactg 4500

5

cgaacttcga	ctcgagcacc	cgtcgcaatg	ggaagcggag	tgggatcgcc	aagcctacct	4560
catcaatcgc	agcgtcaact	gccttctgca	gcgatcagca	caagggttgg	acagcatgtt	4620
ggcaaccgga	atgggtctaca	aggtcttctc	ctccctcgte	gactatgccg	atgggtacaa	4680
gggtctgcag	gaggttgctt	tgcacagcca	agagctcgag	ggcacagcaa	aagtgcgctt	4740
ccaaactccc	tggggagggt	tcgtctgcaa	tcccatgtgg	attgacagct	gtgggtcagac	4800
gaccggcttc	atgatgaact	gtcatcagac	tacgccaat	gactacgtct	acgtcaatca	4860
tggctggaag	tcgatgagat	tggccaaggc	gttccgtgaa	gatgggtacct	atcggactta	4920
tatccggatg	aggcccattg	atagcaccaa	gttcgctggg	gaactgtaca	ttcttgatga	4980
ggatgacact	gtggttggtg	tttatggaga	cataacaatg	agtaaacacga	atctttgtca	5040
acaacaatgc	taacgtaaag	aacagttcca	aggtttgccg	cgacgagttc	tcaacacagt	5100
cctgccatct	gccaacgcgg	ttccagttga	tgctcccatg	gttcaagacg	acctcca	5157

<210> 3

<211> 7339

<212> DNA

<213> Fusarium graminearum

<400> 3

atcgagtttc	gtgttgctgt	tcattggtgta	gcgaacaata	ccatactcct	ccatcaaagg	60
gctgacgagc	ttggcatgga	cctttgtcat	atgggtgctg	tactctttct	cgctaagaga	120
agggtttttg	taggcccgtga	tggtraggca	aacaagggtg	tcaggttgac	cggtgacagt	180
tgcctcatga	gaggtttctt	gggtatcact	atgaggcgac	atgatgcagg	ttggagtttg	240
ttatcgaaat	gctgggtagg	atttggttga	cagagttgca	taaactggag	ctctggatac	300
gcaaggggtc	ataagagtat	ggatacgccg	tcgattagaa	gctgtcagat	ccagtgcacc	360
aatggccccg	atgacgagag	attgctgctc	ggcactgaaa	gtctgcagcg	tgcttggttc	420
ggagttatct	ctacagctgc	accgaatgac	gaacaagccc	gactatccat	gtccgacgag	480
cccggaccga	cgtgatccga	atgctgtcta	acacttacca	agacaatcat	ggggtatctc	540
tgaattatga	ttgtgttgct	caagcagcta	atcccaactc	ttaatcacgt	caagtgccat	600
gcttacctgt	attatcatct	gtcatcaaca	taccaactcc	tatatatagg	ggattgtgct	660
tcattacctc	tccagtaaca	actacttgga	gttcaaacca	catttcaact	tactctttta	720
gttcattcaa	catgacccca	tcaatgatgg	aggatctcgt	ttttggggac	caaagcacac	780
gctttgcccc	tccactgaaa	gacctactcc	tcaaaggcaa	cagtccttac	ttgacacatt	840
ttgttaaaca	agttcacgca	cttcttagaa	gggagatata	atccttgccg	gcagttcaac	900
agaagctttt	cccaaacttt	gccgacattc	aggaaactcgt	ctccaagtca	gattggggca	960
gtggttaacc	tgccttgaca	agcgtcttag	catgctttta	ccatctttgc	agtttcattc	1020
aggagtatca	taaccaagta	cccatagata	tctcctgacc	atccagcttt	tacgatggac	1080
aaggctgtac	ctttccttcg	gagaacagtc	gcattattgg	actttgcgtt	gggtcactcg	1140
ctgctactgc	tgtcagttgc	tccacatcac	tgagtgaatt	ggtatcagct	ggtgtagatg	1200
ctgttcgtgt	ggcattgcac	gtcggactac	gggtatggcg	aactacctcc	cttttcgatg	1260
taccagacag	gccctccgcc	acttggttca	taattgtgcc	cgaggcagta	ctaccaagag	1320
aatctgcgca	agaccgaact	gactcattca	tcattgaaat	ggtaagcaat	tctttgtcct	1380
tggagctctc	ataaaacata	ttcactcttc	tttagggact	tgtctgatca	tcagttcctt	1440
acatcagctc	ggtcgcacat	cacaacatga	ccatcagtg	tccaccatcc	gtcctcgaaa	1500
agttcattca	cagtatatca	acatcaccga	aagattctct	tccagtgcg	atctatgctc	1560
cgtaccacgc	cagccatctt	tacagcatgg	atgatgtaga	cgaggtcctt	agcctgtctg	1620
caacttcttt	tgcacagag	tccatcatte	cactcatttc	aagctcctcg	ggtgacgagt	1680
tacagccact	caagtatgca	gatctactcc	gctgctgtgt	tagtgatatg	ctcatacagc	1740
cactggatct	taccaaggto	tcacaagcag	tggcccagct	tctcgagggt	agctcatcta	1800
cacgtgccat	aataaagcct	atagcaacca	gcgtctccaa	cagtctagt	tctgttttgg	1860
agccgacgct	agcagaacga	tgcgcccgtg	acaacagcat	ggggcccaaa	gcctcgacca	1920
gccactcacc	agcagagaca	caaaccgagt	catcaagcaa	gaactccaaa	attgcgattg	1980

ttgctatgtc	tggctcgctt	ccagacgcag	ctgacttgag	tgaattctgg	gatcttctct	2040
acgaaggtcg	cgatgttcat	cgacaaattc	ccgaggaccg	attcaacgca	gagctccatt	2100
acgacgctac	tgggcgacgt	aagaacacca	gcaaggtcat	gaatggctgc	ttcatcaagg	2160
aaccaggact	gttcgacgct	aggttcttca	acatgtctcc	aaaggaagct	gagcagtcgg	2220
atcctggcca	gcgaatggcc	ctcgaaacag	cttacgagge	gcttgagatg	gctagtatcg	2280
taccagacag	aacaccttcg	acacagagag	atcgtgttgg	tgtgttctac	ggcatgacta	2340
gcatgattg	gagagaggtc	aacagtgggc	agaatgtcga	cacttatctt	attcctggta	2400
agttgtttgt	tgcggtttca	cgacaatact	gattgcagac	tacaggtggc	aacagagcgt	2460
tcactcctgg	tcgactcaac	tacttcttca	agttcagtgg	gcctagcgct	agtgttgata	2520
eggcttgctc	ctccagtctc	gttggcttgc	acttggcttg	taattccctc	tggagaaatg	2580
attgcgatac	agctattgct	ggcggaacca	atgtcatgac	taaccctgac	aacttcgctg	2640
gtttggaccg	aggccacttc	ctatctagaa	ccggttaagt	tctcacagat	cttggaagct	2700
tccttctttt	tagtgttttt	actgacottg	ttccaggcaa	ctgcaacacc	tttgacgatg	2760
gagcagacgg	atactgtcga	gctgatggcg	tcggaaccat	catcctcaag	cggttgagg	2820
acgccgaagc	tgacaatgac	cctattctcg	gtgtcattct	gggcgcttac	acaaaccact	2880
cagccgaagc	agtatccatc	actcgaccac	atgccggagc	tcaagagtac	atcttctcca	2940
aactcctccg	tgagtcgggc	accgatccct	acaacgttag	ctacatcgag	atgcacggca	3000
caggcactca	agccggcgac	gcaaccgaga	tgacatccgt	cctcaagacg	tttgctccta	3060
ccagcggctt	cggcggctga	ttgcctcacc	aaaaccttca	cttgggttca	gtcaaggcca	3120
atgtcgggca	cggtgaatcc	gcctctggta	tcattgtctt	gatcaagacg	ctgcttatga	3180
tggagaagaa	catgatcccc	ccgcattgtg	gtatcaagac	aaagatcaat	caccattttc	3240
ctacggatct	cactcagcgc	aatgtccata	tcgccaaagt	tccgacatct	tggacaagat	3300
cggttcaagc	caatccacgc	attgctttcg	tcaataactt	ctctgccgct	ggtggtaact	3360
ctgctgtcct	actgcaagat	gctcctcagc	catcggtagt	ttcggatgtt	acagaccctc	3420
gcacatccca	tggtgtcact	atgtccgctc	gatcagcaga	ttccctcagg	aagaacctcg	3480
ccaatctcaa	ggagcttgta	gaaggccaag	gtgactcgga	ggtcggcttc	ctgagcaagc	3540
tgtcctacac	aaccaccgcc	aggcgcatgc	atcatcaatt	ccgagcttcg	gtcacagcac	3600
agactcgtga	acagctgctg	aagggccttg	attccgccat	tgaacgccag	gatgtgaaga	3660
ggatccccgc	cgccgcgccc	tctgtcgggt	ttgtgtttag	cgcccaaggc	gcccataacc	3720
gtggtatggg	caaggagtac	tttacatctt	tcacagcctt	ccgctctgag	atcatgtctt	3780
acgacagtat	cgcccaagcc	caaggcttcc	cgtcaatcct	cccactgac	cgaggagagg	3840
tgggaagctga	ctcgttgagt	cctgttgaga	tccagctggg	tctcacttgc	ctgcagatgg	3900
cactggccaa	gctatggaag	tcattcggtg	ttgagccagg	ctttgttctc	ggacacagct	3960
taggccacta	tgctgcttta	cacgtcgctg	gtgttctgtc	cgccaatgat	accatttacc	4020
tcactggcat	cagagcgcag	ctgctcgtgg	ataagtgcc	ggcaggaacc	cactcaatgc	4080
tggcagtga	ggcatcctta	ctacagatcc	aacagttcct	cgatgccaac	attcacgagg	4140
ttgcatgtgt	caatggttca	cgagaagtgc	tcacagtgg	acgcgttgcc	gacattgacc	4200
agctgggttg	cctttttgtc	gctgacaaca	tcaaggcgac	ccgcgtcaag	gtgccatttg	4260
ccttccactc	agcgcaggtt	gaccccatte	tctccgactt	ggatacagcg	gcgtcgcgcg	4320
tcactctcca	ctccctccag	attcctgttc	tttgtgccct	tgacagctct	gtcatcagcc	4380
ctggaaacca	tgggtgtcatt	ggtccccctc	atctacagcg	acattgtcgt	gagacagtca	4440
actttgaggg	tgctctacat	gctgcggagc	acgagaagat	catcaacaag	acatcaactc	4500
tatggatcga	gattgggtccc	catgttgtct	gctctacctt	cctcaagtcc	agccttggtc	4560
caagcacacc	tgctatcgca	tcgcttcgoc	gaaatgacga	ttgctggaag	gtgttggtcg	4620
atggtttgag	cagtctctac	agcagcgggt	tgacaattga	cttgaacgag	tatcatcgcg	4680
acttcaaggc	ctctcaccag	gtacttcgtc	tgccttggtt	cagctggggag	cacaagaatt	4740
actggatata	gtacaagtac	gattgggtcct	tggctaaagg	tgatcctcca	attgccccta	4800
acagctcggg	tgaagcagtc	tcagctttat	caacaccctc	ggtccagaag	attctacagg	4860
agacttccct	tgatcaggta	ttgactatcg	ttgctgagac	agatctcgcg	agccctctat	4920

7

tgtcagaggt tgcccaaggt catcgggtca acggtgtcaa agtctgcaca tcttccgtgt 4980  
 acggtgatgt tggcttgacy ctgggtaagt acattttgga caactaccgc accgacttag 5040  
 agggttatgc ggtcgatgtt caagggtatt aggtccacaa gccacttctt ctcaaagaag 5100  
 acatgaacgg aacgccccag gctacaccgt tccgcacga agtgcgatac ccaatccaga 5160  
 gcaccacggc gctgatgagc atctccacca ctggcccca cggtcagcac atcaagcatg 5220  
 ctaactgcga acttcgactc gagcatccgt cgcaatggga agcggagtgg gatcgccaag 5280  
 cctacctcat caatcgcagc gtcaactgcc ttctgcagcg atcagcacia ggtttggaca 5340  
 gcatggttggc aaccgggaatg gtctacaagg tcttctctc cctcgtcgac tatgccgatg 5400  
 gctacaaggg tctgcaggag gttgtcttgc acagccaaga gctcgagggc acagcaaaag 5460  
 tgggtctcca aactccctcg ggaggtttcg tctgcaatcc catgtggatt gacagctgtg 5520  
 gtcagacgac cggcttcatg atgaactgtc atcagactac gcccaatgac tacgtctacg 5580  
 tcaatcatgg ctggaagtgc atgagattgg ccaaggcggt cctggaagat ggtacctatc 5640  
 ggacttatat cgggatgagg cccattgata gcaccaagtt cgctggtgac ttgtacattc 5700  
 ttgatgagga tgacactgtg gttgggtgtt atggagacat aacaatgagt aacacgaatc 5760  
 tttgtcaaca acaatgctaa cgtaaagaac agttccaagg tttgccgcga cgagttctca 5820  
 acacagtctt gccatctgcc aacggcggtc cagttgatgc tcccatgggt cgacgggacg 5880  
 tgcctccatc aagaatggat gtgcctcccg tcagggtccgg tgaaggggcca cccacttcag 5940  
 caccacgca gcaagctatc gctctgccgt tcgcagccga tacatccatg gactcccgat 6000  
 tgagacctct tcttcgcate ttgtcagaag agatcgggtc cggctctgac gttctttcgg 6060  
 acgatgaact cgactttgcg gaccacgggt tcgactcact cctctcattg accatcactg 6120  
 gtcgcacgag tgaggaattg ggtctcgacg ttgaatctac agcattcatg aactgtccca 6180  
 ctttggggcag ctttaaattg ttcctaggac ttgtcgatca ggacaataag ggcagcagcg 6240  
 gcagtgatgg cagtggtagg agcagtcacg caccgggtac cgagtctggc gctactacac 6300  
 cacctatgag cgaagaggac caggacaaga tagtcagcag tcaactcgct caccagttcc 6360  
 aagccagttc gacgcttcta cagggcagtc ccagtaaagc tcgctcgact ttgttcttgc 6420  
 raccagatgg ctcgggatct gccacatcct acgcttccct tccccgac tctccagacg 6480  
 gagatgttgc tgtctacggg ttgaactgtc catgggtgaa ggactctagt tacctcgtcg 6540  
 agtttggaact caagggttg acagagctct atgtcaacga gatactccgt cgcaagccac 6600  
 agggctcctta caatttggga ggatggtcag cgggtggcat ttgcgcttat gaagctgcc 6660  
 tgatcctcac cagagcagga caccaagtgc atcgcccttat cttgattgac tctcccaatc 6720  
 ccgttggtct tgagaagcta cctcctcgct tgtaogattt cctcaattcg cagaatgtct 6780  
 ttggatcaga caaccgcac agcactgctg gaacaagcgt caaagctcca gaatggcttc 6840  
 ttgcacattt cctggccttc attgacgctc tggatgctta tgtcgcagtg ccttgggact 6900  
 ctggtctagt cggcttagca tcaccgctcc ctgcaccgcc gcagacatac atgctgtggg 6960  
 cagaagacgg agtttgcaaa gactctgata gtgctcgctc cgagtaccgt gacgatgacc 7020  
 cacycgagat gagatggctg ttggagaaca gaacaaactt tgggtccgaat ggttgggagg 7080  
 cgctacttgg tggtaaagag ggtttgttca tggatcggt tgcggaagcg aatcatttta 7140  
 gtatgttgaa gagaggacgg aatgcggaat atgtctctgc attcctggct cgggccttgg 7200  
 acaattagcg atggagaggg ctgcttctgt atttatgtt attcttctt gatgttctga 7260  
 tctcttttgc tttctcatca ttcagataga agcagcagat tgccttcatt ctttataaat 7320  
 cactacagaa ataagtttg 7339

&lt;210&gt; 4

&lt;211&gt; 6608

&lt;212&gt; DNA

<213> *Fusarium graminearum*

&lt;400&gt; 4

atgaccccat caatgatgga ggtattcggt tttggggacc aaagcacacg ctttggccct 60  
 ccactgaaag acctactcct caaaggcaac agtccttact tgacacattt tgttaaacia 120  
 gttcacgcac ttcttagaag ggagatatca tccttgcgag cagttcaaca gaagcttttc 180

```

ccaaactttg cgcacattca ggaactcgtc tccaagtcag attggggcag tggtaaccct 240
gctttgacaa ggcgttttagc atgcttttac catctttgca gtttcattca ggagtatcat 300
aaccaagtac ccatagatct ctctgacca tccagctttt acgatggaca aggtcgtacc 360
cttctctcgg agaacagtcg cattattgga ctttgcggtg gttcactcgc tgcactcgtc 420
gtcagttgct ccacatcact gagtgaattg gtatcagctg gtgtagatgc tgttcgtgtg 480
gca-tgcacg tgggactacg ggtatggcga actacctccc ttttcgatgt accagacagg 540
ccctccgcca cttggttcat aattgtgccc gaggcagtac taccaagaga atctgcgcaa 600
gaccgacttg actcattcat cattgaaatg gtaagcaatt ctttgtcctt ggagctctca 660
taaacacatat tcaactctct ttagggactt gctcgatcat cagttcctta catcagctcg 720
gtcgcacatc acaacatgac catcagtggt ccaccatccg tccctgaaaa gttcattcac 780
agtatatcaa catcaccgaa agattctctt ccagtgcgca tctatgctcc gtaccacgcc 840
agccatcttt acagcagtga tgatgtagac gaggtcctta gcctgtctgc acctctcttt 900
gcatcagagt ccatcattcc actcatttca agctcctcgg gtgacgagtt acagccactc 960
aagratgcag atctactccg ctgctgtggt agtgatatgc tcatacagcc actggatctt 1020
accaaggctt cacaagcagt ggcccagctt ctcgagggtta gctcatctac acgtgccata 1080
ataaagccta tagcaaccag cgtctccaac agtctagtgt ctgttttgga gccgacgcta 1140
gcagaacgat ggcgcgtgga caacagcatg gggcccaaaag cctcgaccag ccactcatca 1200
gcagagacac aaaccgagtc atcaagcaag aactccaaaa ttgcgattgt tgctatgtct 1260
ggtcgctttc cagacgcagc tgacttgagt gaattctggg atcttctcta cgaaggctgc 1320
gatgttcacg gacaaattcc cgaggaccga ttcaacgcag agctccatta cgacgctact 1380
gggcgacgta agaacaccag caaggctcatg aatggctgct tcatcaagga accaggactg 1440
ttcgacgcta ggttcttcaa catgtctcca aaggaagctg agcagtcgga tcctggccag 1500
cgaatggccc tcgaaacagc ttacgaggcg cttgagatgg ctagtatcgt accagacaga 1560
acaccttcga cacagagaga tcgtgttggg gtgttctacg gcatgactag cgatgattgg 1620
agagaggtca acagtgggca gaatgtcgac acttatttta ttcttggtaa gttgtttgtt 1680
cgcgtttcac gacaatactg attgcagact acaggtggca acagagcgtt cactcctggt 1740
cgactcaact acttcttcaa gttcagtggg cctagcgcta gtgttgatac ggcttgctcc 1800
tccagtctcg ttggcttgca cttggcttgt aattccctct ggagaaatga ttgcgataca 1860
gctattgchg gcggaaccaa tgtcatgact aacctgaca acttcgctgg tttggaccga 1920
ggccacttcc tatctagaac cggtaagtct ctacagatc ttggaagctt ccttcttttt 1980
agtgttttta ctgaccttgt tccaggcaac tgcaaacctt ttgacgatgg agcagacgga 2040
tactgtogag ctgatggcgt cggaaccatc atcctcaagc ggcttgagga cgccgaagct 2100
gacaatgacc ctattctcgg tgtcattctg ggcgcttaca caaaccactc agccgaagca 2160
gtatccatca ctcgaccaca tgccggagct caagagtaca tcttctccaa actcctcgt 2220
gagtcgggca ccgatcccta caacgttagc tacatcgaga tgcacggcac aggcactcaa 2280
gccggcgacg caaccgagat gacatccgtc ctcaagacgt ttgctcctac cagcggcttc 2340
ggcggtcgat tgctcacca aaaccttcac ttgggttcag tcaaggccaa tgtcgggcac 2400
ggtgaatccg catctggtat cattgtctct atcaagacgc tgcttatgat ggagaagaac 2460
atgatccgcg cgcattgtgg tatcaagaca aagatcaatc accattttcc tacggatctc 2520
actcagcgca atgtccatat cgccaaagtt ccgacatctt ggacaagatc ggggtcaagcc 2580
aatccacgca ttgctttcgt caataacttc tctgccgctg gtggtaactc tgctgtccta 2640
ctgcaagatg ctctcagcc atcggtagt tccgatgta cagaccctcg cacatcccat 2700
gttgtcacta tgtccgctcg atcagcagat tccctcagga agaacctcgc caatctcaag 2760
gagcttgtag aaggccaagg tgactcggag gtcggcttcc tgagcaagct gtcctacaca 2820
accaccgcca ggcgcatgca tcatcaatc cgagcttcgg tcacagcaca gactcgtgaa 2880
cagctgctga agggccttga ttccgccatt gaacgccagg atgtgaagag gatccccgcc 2940
gccgcgccct ctgtcggctt tgtgttttagc ggccaaggcg cccaataccg tggatatggg 3000
aaggagtact ttacatctct cacagccttc cgctctgaga tcatgtotta cgacagtatc 3060
gcccaagccc aaggcttccc gtcaatcctc ccactgatcc gaggagaggt ggaagctgac 3120

```

togttgagtc	ctggtgagat	ccagctgggt	ctcacttgcc	tgcagatggc	actggccaag	3180
ctatggaagt	cattcggtgt	tgagccagge	tttgttctcg	gacacagctt	aggccactat	3240
gctgctttac	acgtcgctgg	tggtctgtcc	gccaatgata	ccattttacct	cactggcattc	3300
agagcgcagc	tgctcgtgga	taagtgccag	gcaggaaccc	actcaatgct	ggcagtgagg	3360
gcataccttac	tacagatcca	acagttcctc	gatgccaaaca	ttcacgaggt	tgcatgtgtc	3420
aatgggttcac	gagaagtcgt	catcagtgga	cgcgttgccg	acattgacca	gctgggttggc	3480
cttttgctcg	ctgacaacat	caaggcgacc	cgcgtcaagg	tgccatttgc	cttccactca	3540
gcgcaggttg	acccattctt	ctccgacttg	gatacagcgg	cgctcgcgct	caccttccac	3600
tccctccaga	ttcctgttct	ttgtgcccct	gacagctctg	tcatacagccc	tggaaacctat	3660
gggtgtcattg	gtcccccttca	tctacagcga	cattgtcgtg	agacagtcaa	ctttgaggggt	3720
gctctacatg	ctgcggagca	cgagaagatc	atcaacaaga	catcaactct	atggatcgag	3780
actgggtcccc	atgttgtctg	ctctaccttc	ctcaagtcca	gccttggtcc	aagcacacct	3840
gctatcgcac	cgcttcgccc	aaatgacgat	tgctggaagg	tgttggctga	tggtttgagc	3900
agtctctaca	gcagcgggtt	gacaattgac	ttgaacgagt	atcatcgcca	cttcaaggcc	3960
tctcaccagg	tacttcgtct	gccttggtac	agctgggagc	acaagaatta	ctggatacag	4020
tacaagtacg	attgggtcctt	ggctaaagggt	gatcctccaa	ttgcccctaa	cagctcgggtt	4080
gaagcagtcct	cagcttttatc	aacacctctg	gtccagaaga	ttctacagga	gacttccctt	4140
gatcaggtat	tgactatcgt	tgctgagaca	gatctcgcca	gccctctatt	gtcagagggtt	4200
gcccaaggtc	atcgggtcaa	cgggtgtcaa	gtctgcacat	cttccgtgta	cgctgatgtt	4260
ggcttgacgc	tgggtaagta	cattttggac	aactaccgca	ccgacttaga	gggttatgct	4320
gtcgatgttc	acggtattga	ggctccacaag	ccacttcttc	tcaaagaaga	catgaacgga	4380
acgccccagg	ctacaccggt	ccgcacgcaa	gtgcgatacc	caatccagag	caccacggcg	4440
ctgatgagca	tctccaccac	tggccccaac	ggtcagcaca	tcaagcatgc	taactgcgaa	4500
cttcgactcg	agcatccgtc	gcaatgggaa	gcggagtggg	atcgccaagc	ctacctcatc	4560
aatcgcagcg	tcaactgcct	tctgcagcca	tcagcacaag	gtttggacag	catgttggca	4620
acgggaatgg	tctacaagggt	cttctcctcc	ctcgtcgact	atgccgatgg	ctacaagggt	4680
ctgcaggagg	ttgtcttgca	cagccaagag	ctcgaggcca	cagcaaaagt	gcgcttccaa	4740
actccctcgg	gagggtttcgt	ctgcaatccc	atgtggattg	acagctgtgg	tcagacgacc	4800
ggcttcatga	tgaactgtca	tcagactacg	cccaatgact	acgtctacgt	caatcatggc	4860
tggaaagtcga	tgagattggc	caaggcggtc	cgtgaagatg	gtacctatcg	gaacttatatc	4920
cggatgagge	ccattgatag	caccaagttc	gctgggtgact	tgtacattct	tgatgaggat	4980
gacactgtgg	ttgggtgttta	tggagacata	acaatgagta	acacgaatct	ttgtcaacaa	5040
caatgctaac	gtaaagaaca	gttccaagggt	ttgccgcgac	gagttctcaa	cacagtcttg	5100
ccatctgccca	acgcgggttc	agttgatgct	cccatgggtc	gacgggacgt	gcctccatca	5160
agaatggatg	tgcctcccgt	cagggtccgggt	gaagggccac	ccacttcagc	acccacgcag	5220
caagctatcg	ctctgccggt	cgcagccgat	acatccatgg	actcccgaat	gagacctctt	5280
cttcgcactct	tgtcagaaga	gatcggtctc	ggctcttgacg	ttcttttcgga	cgatgaactc	5340
gactttgcgg	accacgggtgt	cgactcactc	ctctcattga	ccatcactgg	tcgcatgcgt	5400
gaggaattgg	gtctcgacgt	tgaatctaca	gcattcatga	actgtcccac	tttgggcagc	5460
tttaaattgt	tcctaggact	tgtcgatcag	gacaataaagg	gcagcagcgg	cagtgatggc	5520
agtggtagga	gcagtcagc	accgggtacc	gagtcctggcg	ctactacacc	acctatgagc	5580
gaagaggacc	aggacaagat	agtcagcagt	cactcgcttc	accagttcca	agccagttcg	5640
acgcttctac	agggcagtc	cagtaaagct	cgtctgactt	tgttcttgct	accagatggc	5700
tcgggatctg	ccacatccta	cgttccctt	cccccgatct	ctccagacgg	agatgttgct	5760
gtctacgggt	tgaactgtcc	atggctgaag	gactctagtt	acctcgtcga	gtttggactc	5820
aagggtcttga	cagagctcta	tgtcaacgag	atactccgtc	gcaagccaca	gggtccttac	5880
aatttgggag	gatgggtcagc	cgggtggcatt	tgcgcttatg	aagctgccct	gatcctcacc	5940
agagcaggac	accaagtcga	tgccttatac	ttgattgact	ctcccaatcc	cgttgggtctt	6000
gagaagctac	ctcctcgtct	gtacgatttc	ctcaattcgc	agaatgtctt	tggatcagac	6060

## 10

aacccgcaca gcactgctgg aacaagcgtc aaagctccag aatggcttct tgcacatttc 6120  
 ctggccttca ttgacgtctt ggatgcttat gtcgcagtgc cttgggactc tggcttagtc 6180  
 ggtctagcat caccgctccc tgcaccgccg cagacataca tgctgtgggc agaagacgga 6240  
 gttcgcaaag actctgatag tgctcgtccc gactaccgtg acgatgaccc acgcgagatg 6300  
 agatggctgt tggagaacag aacaaacttt ggtccgaatg gttgggagggc gctacttggg 6360  
 ggtaagagg gtttgttcat ggatcggatt gcggaagcga atcatttttag tatgttgaag 6420  
 agaggacgga atgcggaata tgtctctgca ttctctggctc gggccttggga caattagcga 6480  
 tggagagggc tgtttgtgta tttatgttta ttcttctttg atgtttgtat ttctttgtct 6540  
 ttcctcatcat tcagatagaa gcagcagatt gccttcattc tttataaatc actacagaaa 6600  
 taagtctg 6608

&lt;210&gt; 5

&lt;211&gt; 6219

&lt;212&gt; DNA

&lt;213&gt; Fusarium graminearum

&lt;220&gt;

&lt;221&gt; CDS

&lt;222&gt; (1)..(6216)

&lt;400&gt; 5

atg acc cca tca atg atg gag gta ttc gtt ttt ggg gac caa agc aca 48  
 Met Thr Pro Ser Met Met Glu Val Phe Val Phe Gly Asp Gln Ser Thr  
 1 5 10 15  
 cgc ttt gcc cct cca ctg aaa gac cta ctc ctc aaa ggc aac agt cct 96  
 Arg Phe Ala Pro Pro Leu Lys Asp Leu Leu Leu Lys Gly Asn Ser Pro  
 20 25 30  
 tac ttg aca cat ttt gtt aaa caa gtt cac gca ctt ctt aga agg gag 144  
 Tyr Leu Thr His Phe Val Lys Gln Val His Ala Leu Leu Arg Arg Glu  
 35 40 45  
 ata tca tcc ttg ccg gca gtt caa cag aag ctt ttc cca aac ttt gcc 192  
 Ile Ser Ser Leu Pro Ala Val Gln Gln Lys Leu Phe Pro Asn Phe Ala  
 50 55 60  
 gac att cag gaa ctc gtc tcc aag tca gat tgg ggc agt ggt aac cct 240  
 Asp Ile Gln Glu Leu Val Ser Lys Ser Asp Trp Gly Ser Gly Asn Pro  
 65 70 75 80  
 gct ttg aca agc gct tta gca tgc ttt tac cat ctt tgc agt ttc att 288  
 Ala Leu Thr Ser Ala Leu Ala Cys Phe Tyr His Leu Cys Ser Phe Ile  
 85 90 95  
 cac ttt tac gat gga caa ggt cgt acc ttt cct tgg gag aac agt cgc 336  
 His Phe Tyr Asp Gly Gln Gly Arg Thr Phe Pro Ser Glu Asn Ser Arg  
 100 105 110  
 att att gga ctt tgc gtt ggt tca ctc gct gct act gct gtc agt tgc 384  
 Ile Ile Gly Leu Cys Val Gly Ser Leu Ala Ala Thr Ala Val Ser Cys  
 115 120 125

## 11

tcc aca tca ctg agt gaa ttg gta tca gct ggt gta gat gct gtt cgt	432
Ser Thr Ser Leu Ser Glu Leu Val Ser Ala Gly Val Asp Ala Val Arg	
130 135 140	
gtg gca ttg cac gtc gga cta cgg gta tgg cga act acc tcc ctt ttc	480
Val Ala Leu His Val Gly Leu Arg Val Trp Arg Thr Thr Ser Leu Phe	
145 150 155 160	
gat gta cca gac agg ccc tcc gcc act tgg ttc ata att gtg ccc gag	528
Asp Val Pro Asp Arg Pro Ser Ala Thr Trp Phe Ile Ile Val Pro Glu	
165 170 175	
gca gta cta cca aga gaa tct gcg caa gac cga ctt gac tca ttc atc	576
Ala Val Leu Pro Arg Glu Ser Ala Gln Asp Arg Leu Asp Ser Phe Ile	
180 185 190	
att gaa atg gga ctt gct cga tca tca gtt cct tac atc agc tcg gtc	624
Ile Glu Met Gly Leu Ala Arg Ser Ser Val Pro Tyr Ile Ser Ser Val	
195 200 205	
gca cat cac aac atg acc atc agt ggt cca cca tcc gtc ctc gaa aag	672
Ala His His Asn Met Thr Ile Ser Gly Pro Pro Ser Val Leu Glu Lys	
210 215 220	
ttc att cac agt ata tca aca tca cgg aaa gat tct ctt cca gtg ccg	720
Phe Ile His Ser Ile Ser Thr Ser Pro Lys Asp Ser Leu Pro Val Pro	
225 230 235 240	
atc tat gct ccg tac cac gcc agc cat ctt tac agc atg gat gat gta	768
Ile Tyr Ala Pro Tyr His Ala Ser His Leu Tyr Ser Met Asp Asp Val	
245 250 255	
gac gag gtc ctt agc ctg tct gca cct tct ttt gca tca gag tcc atc	816
Asp Glu Val Leu Ser Leu Ser Ala Pro Ser Phe Ala Ser Glu Ser Ile	
260 265 270	
att cca ctc att tca agc tcc tcg ggt gac gag tta cag cca ctc aag	864
Ile Pro Leu Ile Ser Ser Ser Gly Asp Glu Leu Gln Pro Leu Lys	
275 280 285	
tat gca gat cta ctc cgc tgc tgt gtt agt gat atg ctc ata cag cca	912
Tyr Ala Asp Leu Leu Arg Cys Cys Val Ser Asp Met Leu Ile Gln Pro	
290 295 300	
ctg gat ctt acc aag gtc tca caa gca gtg gcc cag ctt ctc gag gtt	960
Leu Asp Leu Thr Lys Val Ser Gln Ala Val Ala Gln Leu Leu Glu Val	
305 310 315 320	
agc tca tct aca cgt gcc ata ata aag cct ata gca acc agc gtc tcc	1008
Ser Ser Ser Thr Arg Ala Ile Ile Lys Pro Ile Ala Thr Ser Val Ser	
325 330 335	
aac agt cta gtg tct gtt ttg gag ccg acg cta gca gaa cga tgc gcc	1056
Asn Ser Leu Val Ser Val Leu Glu Pro Thr Leu Ala Glu Arg Cys Ala	
340 345 350	

12

gtg gac aac agc atg ggg ccc aaa gcc tgc acc agc cac tca tca gca	1104
Val Asp Asn Ser Met Gly Pro Lys Ala Ser Thr Ser His Ser Ser Ala	
355 360 365	
gag aca caa acc gag tca tca agc aag aac tcc aaa att gcg att gtt	1152
Glu Thr Gln Thr Glu Ser Ser Ser Lys Asn Ser Lys Ile Ala Ile Val	
370 375 380	
gct atg tct ggt cgc ttt cca gac gca gct gac ttg agt gaa ttc tgg	1200
Ala Met Ser Gly Arg Phe Pro Asp Ala Ala Asp Leu Ser Glu Phe Trp	
385 390 395 400	
gat ctt ctc tac gaa ggt cgc gat gtt cat cga caa att ccc gag gac	1248
Asp Leu Leu Tyr Glu Gly Arg Asp Val His Arg Gln Ile Pro Glu Asp	
405 410 415	
cga ttc aac gca gag ctc cat tac gac gct act ggg cga cgt aag aac	1296
Arg Phe Asn Ala Glu Leu His Tyr Asp Ala Thr Gly Arg Arg Lys Asn	
420 425 430	
acc agc aag gtc atg aat ggc tgc ttc atc aag gaa cca gga ctg ttc	1344
Thr Ser Lys Val Met Asn Gly Cys Phe Ile Lys Glu Pro Gly Leu Phe	
435 440 445	
gac gct agg ttc ttc aac atg tct cca aag gaa gct gag cag tgc gat	1392
Asp Ala Arg Phe Phe Asn Met Ser Pro Lys Glu Ala Glu Gln Ser Asp	
450 455 460	
cct ggc cag cga atg gcc ctc gaa aca gct tac gag gcg ctt gag atg	1440
Pro Gly Gln Arg Met Ala Leu Glu Thr Ala Tyr Glu Ala Leu Glu Met	
465 470 475 480	
gct agt atc gta cca gac aga aca cct tgc aca cag aga gat cgt gtt	1488
Ala Ser Ile Val Pro Asp Arg Thr Pro Ser Thr Gln Arg Asp Arg Val	
485 490 495	
ggt gtg ttc tac ggc atg act agc gat gat tgg aga gag gtc aac agt	1536
Gly Val Phe Tyr Gly Met Thr Ser Asp Asp Trp Arg Glu Val Asn Ser	
500 505 510	
ggg cag aat gtc gac act tat ttt att cct ggt ggc aac aga gcg ttc	1584
Gly Gln Asn Val Asp Thr Tyr Phe Ile Pro Gly Gly Asn Arg Ala Phe	
515 520 525	
act cct ggt cga ctc aac tac ttc ttc aag ttc agt ggg cct agc gct	1632
Thr Pro Gly Arg Leu Asn Tyr Phe Phe Lys Phe Ser Gly Pro Ser Ala	
530 535 540	
agt gtt gat acg gct tgc tcc tcc agt ctc gtt ggc ttg cac ttg gct	1680
Ser Val Asp Thr Ala Cys Ser Ser Ser Leu Val Gly Leu His Leu Ala	
545 550 555 560	
tgt aat tcc ctc tgg aga aat gat tgc gat aca gct att gcg ggc gga	1728
Cys Asn Ser Leu Trp Arg Asn Asp Cys Asp Thr Ala Ile Ala Gly Gly	
565 570 575	

13

acc aat gtc atg act aac cct gac aac ttc gct ggt ttg gac cga ggc	1776
Thr Asn Val Met Thr Asn Pro Asp Asn Phe Ala Gly Leu Asp Arg Gly	
580 585 590	
cac ttc cta tct aga acc ggc aac tgc aac acc ttt gac gat gga gca	1824
His Phe Leu Ser Arg Thr Gly Asn Cys Asn Thr Phe Asp Asp Gly Ala	
595 600 605	
gac gga tac tgt cga gct gat ggc gtc gga acc atc atc ctc aag cgg	1872
Asp Gly Tyr Cys Arg Ala Asp Gly Val Gly Thr Ile Ile Leu Lys Arg	
610 615 620	
ctt gag gac gcc gaa gct gac aat gac cct att ctc ggt gtc att ctg	1920
Leu Glu Asp Ala Glu Ala Asp Asn Asp Pro Ile Leu Gly Val Ile Leu	
625 630 635 640	
ggc gct tac aca aac cac tca gcc gaa gca gta tcc atc act cga cca	1968
Gly Ala Tyr Thr Asn His Ser Ala Glu Ala Val Ser Ile Thr Arg Pro	
645 650 655	
cat gcc gga gct caa gag tac atc ttc tcc aaa ctc ctc cgt gag tcg	2016
His Ala Gly Ala Gln Glu Tyr Ile Phe Ser Lys Leu Leu Arg Glu Ser	
660 665 670	
ggc acc gat ccc tac aac gtt agc tac atc gag atg cac ggc aca ggc	2064
Gly Thr Asp Pro Tyr Asn Val Ser Tyr Ile Glu Met His Gly Thr Gly	
675 680 685	
act caa gcc ggc gac gca acc gag atg aca tcc gtc ctc aag acg ttt	2112
Thr Gln Ala Gly Asp Ala Thr Glu Met Thr Ser Val Leu Lys Thr Phe	
690 695 700	
gct cct acc agc ggc ttc ggc ggt cga ttg cct cac caa aac ctt cac	2160
Ala Pro Thr Ser Gly Phe Gly Gly Arg Leu Pro His Gln Asn Leu His	
705 710 715 720	
ttg ggt tca gtc aag gcc aat gtc ggg cac ggt gaa tcc gca tct ggt	2208
Leu Gly Ser Val Lys Ala Asn Val Gly His Gly Glu Ser Ala Ser Gly	
725 730 735	
atc att gct ctg atc aag acg ctg ctt atg atg gag aag aac atg atc	2256
Ile Ile Ala Leu Ile Lys Thr Leu Leu Met Met Glu Lys Asn Met Ile	
740 745 750	
ccg ccg cat tgt ggt atc aag aca aag atc aat cac cat ttt cct acg	2304
Pro Pro His Cys Gly Ile Lys Thr Lys Ile Asn His His Phe Pro Thr	
755 760 765	
gat ctc act cag cgc aat gtc cat atc gcc aaa gtt ccg aca tct tgg	2352
Asp Leu Thr Gln Arg Asn Val His Ile Ala Lys Val Pro Thr Ser Trp	
770 775 780	
aca aga tcg ggt caa gcc aat coa cgc att gct ttc gtc aat aac ttc	2400
Thr Arg Ser Gly Gln Ala Asn Pro Arg Ile Ala Phe Val Asn Asn Phe	
785 790 795 800	

## 14

tct gcc gct ggt ggt aac tct gct gtc cta ctg caa gat gct cct cag	2448
Ser Ala Ala Gly Gly Asn Ser Ala Val Leu Leu Gln Asp Ala Pro Gln	
805 810 815	
cca tcg gta gtt tcg gat gtt aca gac cct cgc aca tcc cat gtt gtc	2496
Pro Ser Val Val Ser Asp Val Thr Asp Pro Arg Thr Ser His Val Val	
820 825 830	
act atg tcc gct cga tca gca gat tcc ctc agg aag aac ctc gcc aat	2544
Thr Met Ser Ala Arg Ser Ala Asp Ser Leu Arg Lys Asn Leu Ala Asn	
835 840 845	
ctc aag gag ctt gta gaa ggc caa ggt gac tcg gag gtc ggc ttc ctg	2592
Leu Lys Glu Leu Val Glu Gly Gln Gly Asp Ser Glu Val Gly Phe Leu	
850 855 860	
agc aag ctg tcc tac aca acc acc gcc agg cgc atg cat cat caa ttc	2640
Ser Lys Leu Ser Tyr Thr Thr Thr Ala Arg Arg Met His His Gln Phe	
865 870 875 880	
cga gct tcg gtc aca gca cag act cgt gaa cag ctg ctg aag ggc ctt	2688
Arg Ala Ser Val Thr Ala Gln Thr Arg Glu Gln Leu Leu Lys Gly Leu	
885 890 895	
gat tcc gcc att gaa cgc cag gat gtg aag agg atc ccc gcc gcc gcg	2736
Asp Ser Ala Ile Glu Arg Gln Asp Val Lys Arg Ile Pro Ala Ala Ala	
900 905 910	
ccc tct gtc ggc ttt gtg ttt agc ggc caa ggc gcc caa tac cgt ggt	2784
Pro Ser Val Gly Phe Val Phe Ser Gly Gln Gly Ala Gln Tyr Arg Gly	
915 920 925	
atg ggc aag gag tac ttt aca tct ttc aca gcc ttc cgc tct gag atc	2832
Met Gly Lys Glu Tyr Phe Thr Ser Phe Thr Ala Phe Arg Ser Glu Ile	
930 935 940	
atg tct tac gac agt atc gcc caa gcc caa ggc ttc cgc tca atc ctc	2880
Met Ser Tyr Asp Ser Ile Ala Gln Ala Gln Gly Phe Pro Ser Ile Leu	
945 950 955 960	
cca ctg atc cga gga gag gtg gaa gct gac tcg ttg agt cct gtt gag	2928
Pro Leu Ile Arg Gly Glu Val Glu Ala Asp Ser Leu Ser Pro Val Glu	
965 970 975	
atc cag ctg ggt ctc act tgc ctg cag atg gca ctg gcc aag cta tgg	2976
Ile Gln Leu Gly Leu Thr Cys Leu Gln Met Ala Leu Ala Lys Leu Trp	
980 985 990	
aag tca ttc ggt gtt gag cca ggc ttt gtt ctc gga cac agc tta ggc	3024
Lys Ser Phe Gly Val Glu Pro Gly Phe Val Leu Gly His Ser Leu Gly	
995 1000 1005	
cac tat gct gct tta cac gtc gct ggt gtt ctg tcc gcc aat gat acc	3072
His Tyr Ala Ala Leu His Val Ala Gly Val Leu Ser Ala Asn Asp Thr	
1010 1015 1020	

15

att tac ctc act ggc atc aga gcg cag ctg ctc gtg gat aag tgc cag	3120
Ile Tyr Leu Thr Gly Ile Arg Ala Gln Leu Leu Val Asp Lys Cys Gln	
1025 1030 1035 1040	
gca gga acc cac tca atg ctg gca gtg agg gca tcc tta cta cag atc	3168
Ala Gly Thr His Ser Met Leu Ala Val Arg Ala Ser Leu Leu Gln Ile	
1045 1050 1055	
caa cag ttc ctc gat gcc aac att cac gag gtt gca tgt gtc aat ggt	3216
Gln Gln Phe Leu Asp Ala Asn Ile His Glu Val Ala Cys Val Asn Gly	
1060 1065 1070	
tca cga gaa gtc gtc atc agt gga cgc gtt gcc gac att gac cag ctg	3264
Ser Arg Glu Val Val Ile Ser Gly Arg Val Ala Asp Ile Asp Gln Leu	
1075 1080 1085	
gtt ggc ctt ttg tgg gct gac aac atc aag gcg acc cgc gtc aag gtg	3312
Val Gly Leu Leu Ser Ala Asp Asn Ile Lys Ala Thr Arg Val Lys Val	
1090 1095 1100	
cca ttt gcc ttc cac tca gcg cag gtt gac ccc att ctc tcc gac ttg	3360
Phe Phe Ala Phe His Ser Ala Gln Val Asp Pro Ile Leu Ser Asp Leu	
1105 1110 1115 1120	
gat aca gcg gcg tgg cgc gtc acc ttc cac tcc ctc cag att cct gtt	3408
Asp Thr Ala Ala Ser Arg Val Thr Phe His Ser Leu Gln Ile Pro Val	
1125 1130 1135	
ctt tgt gcc ctt gac agc tct gtc atc agc cct gga aac cat ggt gtc	3456
Leu Cys Ala Leu Asp Ser Ser Val Ile Ser Pro Gly Asn His Gly Val	
1140 1145 1150	
att ggt ccc ctt cat cta cag cga cat tgt cgt gag aca gtc aac ttt	3504
Ile Gly Pro Leu His Leu Gln Arg His Cys Arg Glu Thr Val Asn Phe	
1155 1160 1165	
gag ggt gct cta cat gct gcg gag cac gag aag atc atc aac aag aca	3552
Glu Gly Ala Leu His Ala Ala Glu His Glu Lys Ile Ile Asn Lys Thr	
1170 1175 1180	
tca act cta tgg atc gag att ggt ccc cat gtt gtc tgc tct acc ttc	3600
Ser Thr Leu Trp Ile Glu Ile Gly Pro His Val Val Cys Ser Thr Phe	
1185 1190 1195 1200	
ctc aag tcc agc ctt ggt cca agc aca cct gct atc gca tgg ctt cgc	3648
Leu Lys Ser Ser Leu Gly Pro Ser Thr Pro Ala Ile Ala Ser Leu Arg	
1205 1210 1215	
cga aat gac gat tgc tgg aag gtg ttg gct gat ggt ttg agc agt ctc	3696
Arg Asn Asp Asp Cys Trp Lys Val Leu Ala Asp Gly Leu Ser Ser Leu	
1220 1225 1230	
tac agc agc ggg ttg aca att gac ttg aac gag tat cat cgc gac ttc	3744
Tyr Ser Ser Gly Leu Thr Ile Asp Leu Asn Glu Tyr His Arg Asp Phe	
1235 1240 1245	

## 16

aag gcc tct cac cag gta ctt cgt ctg cct tgt tac agc tgg gag cac	3792
Lys Ala Ser His Gln Val Leu Arg Leu Pro Cys Tyr Ser Trp Glu His	
1250 1255 1260	
aag aat tac tgg ata cag tac aag tac gat tgg tcc ttg gct aaa ggt	3840
Lys Asn Tyr Trp Ile Gln Tyr Lys Tyr Asp Trp Ser Leu Ala Lys Gly	
1265 1270 1275 1280	
gat cct cca att gcc cct aac agc tcg gtt gaa gca gtc tca gct tta	3888
Asp Pro Pro Ile Ala Pro Asn Ser Ser Val Glu Ala Val Ser Ala Leu	
1285 1290 1295	
tca aca ccc tcg gtc cag aag att cta cag gag act tcc ctt gat cag	3936
Ser Thr Pro Ser Val Gln Lys Tle Leu Gln Glu Thr Ser Leu Asp Gln	
1300 1305 1310	
gta ttg act atc gtt gct gag aca gat ctc gcg agc cct cta ttg tca	3984
Val Leu Thr Ile Val Ala Glu Thr Asp Leu Ala Ser Pro Leu Leu Ser	
1315 1320 1325	
gag gtt gcc caa ggt cat cgg gtc aac ggt gtc aaa gtc tgc aca tct	4032
Glu Val Ala Gln Gly His Arg Val Asn Gly Val Lys Val Cys Thr Ser	
1330 1335 1340	
tcc gtg tac gct gat gtt ggc ttg acg ctg ggt aag tac att ttg gac	4080
Ser Val Tyr Ala Asp Val Gly Leu Thr Leu Gly Lys Tyr Ile Leu Asp	
1345 1350 1355 1360	
aac tac cgc acc gac tta gag ggt tat gcg gtc gat gtt cac ggt att	4128
Asn Tyr Arg Thr Asp Leu Glu Gly Tyr Ala Val Asp Val His Gly Ile	
1365 1370 1375	
gag gtc cac aag cca ctt ctt ctc aaa gaa gac atg aac gga acg ccc	4176
Glu Val His Lys Pro Leu Leu Leu Lys Glu Asp Met Asn Gly Thr Pro	
1380 1385 1390	
cag gct aca ccg ttc cgc atc gaa gtg cga tac cca atc cag agc acc	4224
Gln Ala Thr Pro Phe Arg Ile Glu Val Arg Tyr Pro Ile Gln Ser Thr	
1395 1400 1405	
acg gcg ctg atg agc atc tcc acc act ggc ccc aac ggt cag cac atc	4272
Thr Ala Leu Met Ser Ile Ser Thr Thr Gly Pro Asn Gly Gln His Ile	
1410 1415 1420	
aag cat gct aac tgc gaa ctt cga ctc gag cat ccg tcg caa tgg gaa	4320
Lys His Ala Asn Cys Glu Leu Arg Leu Glu His Pro Ser Gln Trp Glu	
1425 1430 1435 1440	
gcg gag tgg gat cgc caa gcc tac ctc atc aat cgc agc gtc aac tgc	4368
Ala Glu Trp Asp Arg Gln Ala Tyr Leu Ile Asn Arg Ser Val Asn Cys	
1445 1450 1455	
ctt ctg cag cga tca gca caa ggt ttg gac agc atg ttg gca acc gga	4416
Leu Leu Gln Arg Ser Ala Gln Gly Leu Asp Ser Met Leu Ala Thr Gly	
1460 1465 1470	

17

atg gtc tac aag gtc ttc tcc tcc ctc gtc gac tat gcc gat ggc tac	4464
Met. Val Tyr Lys Val Phe Ser Ser Leu Val Asp Tyr Ala Asp Gly Tyr	
1475 1480 1485	
aag ggt ctg cag gag gtt gtc ttg cac agc caa gag ctc gag ggc aca	4512
Lys Gly Leu Gln Glu Val Val Leu His Ser Gln Glu Leu Glu Gly Thr	
1490 1495 1500	
gca aaa gtg cgc ttc caa act ccc tcg gga ggt ttc gtc tgc aat ccc	4560
Ala Lys Val Arg Phe Gln Thr Pro Ser Gly Gly Phe Val Cys Asn Pro	
1505 1510 1515 1520	
atg tgg att gac agc tgt ggt cag acg acc ggc ttc atg atg aac tgt	4608
Met. Trp Ile Asp Ser Cys Gly Gln Thr Thr Gly Phe Met Met Asn Cys	
1525 1530 1535	
cat cag act acg ccc aat gac tac gtc tac gtc aat cat ggc tgg aag	4656
His Gln Thr Thr Pro Asn Asp Tyr Val Tyr Val Asn His Gly Trp Lys	
1540 1545 1550	
tcg atg aga ttg gcc aag gcg ttc cgt gaa gat ggt acc tat cgg act	4704
Ser Met Arg Leu Ala Lys Ala Phe Arg Glu Asp Gly Thr Tyr Arg Thr	
1555 1560 1565	
tat atc cgg atg agg ccc att gat agc acc aag ttc gct ggt gac ttg	4752
Tyr Ile Arg Met Arg Pro Ile Asp Ser Thr Lys Phe Ala Gly Asp Leu	
1570 1575 1580	
tac att ctt gat gag gat gac act gtg gtt ggt gtt tat gga gac ata	4800
Tyr Ile Leu Asp Glu Asp Asp Thr Val Val Gly Val Tyr Gly Asp Ile	
1585 1590 1595 1600	
aca ttc caa ggt ttg ccg cga cga gtt ctc aac aca gtc ttg cca tct	4848
Thr. Phe Gln Gly Leu Pro Arg Arg Val Leu Asn Thr Val Leu Pro Ser	
1605 1610 1615	
gcc aac gcg gtt cca gtt gat gct ccc atg ggt cga cgg gac gtg cct	4896
Ala Asn Ala Val Pro Val Asp Ala Pro Met Gly Arg Arg Asp Val Pro	
1620 1625 1630	
cca tca aga atg gat gtg cct ccc gtc agg tcc ggt gaa ggg cca ccc	4944
Pro Ser Arg Met Asp Val Pro Pro Val Arg Ser Gly Glu Gly Pro Pro	
1635 1640 1645	
act tca gca ccc acg cag caa gct atc gct ctg ccg ttc gca gcc gat	4992
Thr Ser Ala Pro Thr Gln Gln Ala Ile Ala Leu Pro Phe Ala Ala Asp	
1650 1655 1660	
aca tcc atg gac tcc cga ttg aga cct ctt ctt cgc atc ttg tca gaa	5040
Thr Ser Met Asp Ser Arg Leu Arg Pro Leu Leu Arg Ile Leu Ser Glu	
1665 1670 1675 1680	
gag atc ggt ctc ggt ctt gac gtt ctt tcg gac gat gaa ctc gac ttt	5088
Glu Ile Gly Leu Gly Leu Asp Val Leu Ser Asp Asp Glu Leu Asp Phe	
1685 1690 1695	

## 18

gcg gac cac ggt gtc gac tca ctc ctc tca ttg acc atc act ggt cgc	5136
Ala Asp His Gly Val Asp Ser Leu Leu Ser Leu Thr Ile Thr Gly Arg	
1700 1705 1710	
atg cgt gag gaa ttg ggt ctc gac gtt gaa tct aca gca ttc atg aac	5184
Met Arg Glu Glu Leu Gly Leu Asp Val Glu Ser Thr Ala Phe Met Asn	
1715 1720 1725	
tgt ccc act ttg ggc agc ttt aaa ttg ttc cta gga ctt gtc gat cag	5232
Cys Pro Thr Leu Gly Ser Phe Lys Leu Phe Leu Gly Leu Val Asp Gln	
1730 1735 1740	
gac aat aag ggc agc agc ggc agt gat ggc agt ggt agg agc agt cca	5280
Asp Asn Lys Gly Ser Ser Gly Ser Asp Gly Ser Gly Arg Ser Ser Pro	
1745 1750 1755 1760	
gca ccg ggt acc gag tct ggc gct act aca cca cct atg agc gaa gag	5328
Ala Pro Gly Thr Glu Ser Gly Ala Thr Thr Pro Pro Met Ser Glu Glu	
1765 1770 1775	
gac cag gac aag ata gtc agc agt cac tcg ctt cac cag ttc caa gcc	5376
Asp Gln Asp Lys Ile Val Ser Ser His Ser Leu His Gln Phe Gln Ala	
1780 1785 1790	
agt tcg acg ctt cta cag ggc agt ccc agt aaa gct cgc tcg act ttg	5424
Ser Ser Thr Leu Leu Gln Gly Ser Pro Ser Lys Ala Arg Ser Thr Leu	
1795 1800 1805	
ttc ttg cta cca gat ggc tcg gga tct gcc aca tcc tac gct tcc ctt	5472
Phe Leu Leu Pro Asp Gly Ser Gly Ser Ala Thr Ser Tyr Ala Ser Leu	
1810 1815 1820	
ccc ccg atc tct cca gac gga gat gtt gct gtc tac ggg ttg aac tgt	5520
Pro Pro Ile Ser Pro Asp Gly Asp Val Ala Val Tyr Gly Leu Asn Cys	
1825 1830 1835 1840	
cca tgg ctg aag gac tct agt tac ctc gtc gag ttt gga ctc aag ggc	5568
Pro Trp Leu Lys Asp Ser Ser Tyr Leu Val Glu Phe Gly Leu Lys Gly	
1845 1850 1855	
ttg aca gag ctc tat gtc aac gag ata ctc cgt cgc aag cca cag ggt	5616
Leu Thr Glu Leu Tyr Val Asn Glu Ile Leu Arg Arg Lys Pro Gln Gly	
1860 1865 1870	
cct tac aat ttg gga gga tgg tca gcc ggt ggc att tgc gct tat gaa	5664
Pro Tyr Asn Leu Gly Gly Trp Ser Ala Gly Gly Ile Cys Ala Tyr Glu	
1875 1880 1885	
gct gcc ctg atc ctc acc aga gca gga cac caa gtc gat cgc ctt atc	5712
Ala Ala Leu Ile Leu Thr Arg Ala Gly His Gln Val Asp Arg Leu Ile	
1890 1895 1900	
ttg att gac tct ccc aat ccc gtt ggt ctt gag aag cta cct cct cgc	5760
Leu Ile Asp Ser Pro Asn Pro Val Gly Leu Glu Lys Leu Pro Pro Arg	
1905 1910 1915 1920	

## 19

ttg tac gat ttc ctc aat tgc cag aat gtc ttt gga tca gac aac ccg 5808  
 Leu Tyr Asp Phe Leu Asn Ser Gln Asn Val Phe Gly Ser Asp Asn Pro  
 1925 1930 1935

cac agc act gct gga aca agc gtc aaa gct cca gaa tgg ctt ctt gca 5856  
 His Ser Thr Ala Gly Thr Ser Val Lys Ala Pro Glu Trp Leu Leu Ala  
 1940 1945 1950

cat ttc ctg gcc ttc att gac gct ctg gat gct tat gtc gca gtg cct 5904  
 His Phe Leu Ala Phe Ile Asp Ala Leu Asp Ala Tyr Val Ala Val Pro  
 1955 1960 1965

tgg gac tct ggt cta gtc ggt cta gca tca ccg ctc cct gca ccg ccg 5952  
 Trp Asp Ser Gly Leu Val Gly Leu Ala Ser Pro Leu Pro Ala Pro Pro  
 1970 1975 1980

cag aca tac atg ctg tgg gca gaa gac gga gtt tgc aaa gac tct gat 6000  
 Gln Thr Tyr Met Leu Trp Ala Glu Asp Gly Val Cys Lys Asp Ser Asp  
 1985 1990 1995 2000

agt gct cgt ccc gag tac cgt gac gat gac cca cgc gag atg aga tgg 6048  
 Ser Ala Arg Pro Glu Tyr Arg Asp Asp Asp Pro Arg Glu Met Arg Trp  
 2005 2010 2015

ctg ttg gag aac aga aca aac ttt ggt ccg aat ggt tgg gag gcg cta 6096  
 Leu Leu Glu Asn Arg Thr Asn Phe Gly Pro Asn Gly Trp Glu Ala Leu  
 2020 2025 2030

ctt ggt ggt aaa gag ggt ttg ttc atg gat cgg att gcg gaa gcg aat 6144  
 Leu Gly Gly Lys Glu Gly Leu Phe Met Asp Arg Ile Ala Glu Ala Asn  
 2035 2040 2045

cat ttc agt atg ttg aag aga gga cgg aat gcg gaa tat gtc tct gca 6192  
 His Phe Ser Met Leu Lys Arg Gly Arg Asn Ala Glu Tyr Val Ser Ala  
 2050 2055 2060

ttc ctg gct cgg gcc ttg gac aat tag 6219  
 Phe Leu Ala Arg Ala Leu Asp Asn  
 2065 2070

<210> 6

<211> 2072

<212> PRT

<213> Fusarium graminearum

<400> 6

Met Thr Pro Ser Met Met Glu Val Phe Val Phe Gly Asp Gln Ser Thr  
 1 5 10 15

Arg Phe Ala Pro Pro Leu Lys Asp Leu Leu Leu Lys Gly Asn Ser Pro  
 20 25 30

Tyr Leu Thr His Phe Val Lys Gln Val His Ala Leu Leu Arg Arg Glu  
 35 40 45

20

Ile Ser Ser Leu Pro Ala Val Gln Gln Lys Leu Phe Pro Asn Phe Ala		
50	55	60
Asp Ile Gln Glu Leu Val Ser Lys Ser Asp Trp Gly Ser Gly Asn Pro		
65	70	75 80
Ala Leu Thr Ser Ala Leu Ala Cys Phe Tyr His Leu Cys Ser Phe Ile		
85	90	95
His Phe Tyr Asp Gly Gln Gly Arg Thr Phe Pro Ser Glu Asn Ser Arg		
100	105	110
Ile Ile Gly Leu Cys Val Gly Ser Leu Ala Ala Thr Ala Val Ser Cys		
115 .	120 .	125 .
Ser Thr Ser Leu Ser Glu Leu Val Ser Ala Gly Val Asp Ala Val Arg		
130	135	140
Val Ala Leu His Val Gly Leu Arg Val Trp Arg Thr Thr Ser Leu Phe		
145	150	155 160
Asp Val Pro Asp Arg Pro Ser Ala Thr Trp Phe Ile Ile Val Pro Glu		
165	170	175
Ala Val Leu Pro Arg Glu Ser Ala Gln Asp Arg Leu Asp Ser Phe Ile		
180	185	190
Ile Glu Met Gly Leu Ala Arg Ser Ser Val Pro Tyr Ile Ser Ser Val		
195	200	205
Ala His His Asn Met Thr Ile Ser Gly Pro Pro Ser Val Leu Glu Lys		
210	215	220
Phe Ile His Ser Ile Ser Thr Ser Pro Lys Asp Ser Leu Pro Val Pro		
225	230	235 240
Ile Tyr Ala Pro Tyr His Ala Ser His Leu Tyr Ser Met Asp Asp Val		
245	250	255
Asp Glu Val Leu Ser Leu Ser Ala Pro Ser Phe Ala Ser Glu Ser Ile		
260	265	270
Ile Pro Leu Ile Ser Ser Ser Ser Gly Asp Glu Leu Gln Pro Leu Lys		
275	280	285
Tyr Ala Asp Leu Leu Arg Cys Cys Val Ser Asp Met Leu Ile Gln Pro		
290	295	300
Leu Asp Leu Thr Lys Val Ser Gln Ala Val Ala Gln Leu Leu Glu Val		
305	310	315 320
Ser Ser Ser Thr Arg Ala Ile Ile Lys Pro Ile Ala Thr Ser Val Ser		
325	330	335
Asn Ser Leu Val Ser Val Leu Glu Pro Thr Leu Ala Glu Arg Cys Ala		
340	345	350

## 21

Val Asp Asn Ser Met Gly Pro Lys Ala Ser Thr Ser His Ser Ser Ala  
355 360 365  
Glu Thr Gln Thr Glu Ser Ser Ser Lys Asn Ser Lys Ile Ala Ile Val  
370 375 380  
Ala Met Ser Gly Arg Phe Pro Asp Ala Ala Asp Leu Ser Glu Phe Trp  
385 390 395 400  
Asp Leu Leu Tyr Glu Gly Arg Asp Val His Arg Gln Ile Pro Glu Asp  
405 410 415  
Arg Phe Asn Ala Glu Leu His Tyr Asp Ala Thr Gly Arg Arg Lys Asn  
420 425 430  
Thr Ser Lys Val Met Asn Gly Cys Phe Ile Lys Glu Pro Gly Leu Phe  
435 440 445  
Asp Ala Arg Phe Phe Asn Met Ser Pro Lys Glu Ala Glu Gln Ser Asp  
450 455 460  
Pro Gly Gln Arg Met Ala Leu Glu Thr Ala Tyr Glu Ala Leu Glu Met  
465 470 475 480  
Ala Ser Ile Val Pro Asp Arg Thr Pro Ser Thr Gln Arg Asp Arg Val  
485 490 495  
Gly Val Phe Tyr Gly Met Thr Ser Asp Asp Trp Arg Glu Val Asn Ser  
500 505 510  
Gly Gln Asn Val Asp Thr Tyr Phe Ile Pro Gly Gly Asn Arg Ala Phe  
515 520 525  
Thr Pro Gly Arg Leu Asn Tyr Phe Phe Lys Phe Ser Gly Pro Ser Ala  
530 535 540  
Ser Val Asp Thr Ala Cys Ser Ser Ser Leu Val Gly Leu His Leu Ala  
545 550 555 560  
Cys Asn Ser Leu Trp Arg Asn Asp Cys Asp Thr Ala Ile Ala Gly Gly  
565 570 575  
Thr Asn Val Met Thr Asn Pro Asp Asn Phe Ala Gly Leu Asp Arg Gly  
580 585 590  
His Phe Leu Ser Arg Thr Gly Asn Cys Asn Thr Phe Asp Asp Gly Ala  
595 600 605  
Asp Gly Tyr Cys Arg Ala Asp Gly Val Gly Thr Ile Ile Leu Lys Arg  
610 615 620  
Leu Glu Asp Ala Glu Ala Asp Asn Asp Pro Ile Leu Gly Val Ile Leu  
625 630 635 640  
Gly Ala Tyr Thr Asn His Ser Ala Glu Ala Val Ser Ile Thr Arg Pro  
645 650 655

22

His	Ala	Gly	Ala	Gln	Glu	Tyr	Ile	Phe	Ser	Lys	Leu	Leu	Arg	Glu	Ser	660	665	670
Gly	Thr	Asp	Pro	Tyr	Asn	Val	Ser	Tyr	Ile	Glu	Met	His	Gly	Thr	Gly	675	680	685
Thr	Gln	Ala	Gly	Asp	Ala	Thr	Glu	Met	Thr	Ser	Val	Leu	Lys	Thr	Phe	690	695	700
Ala	Pro	Thr	Ser	Gly	Phe	Gly	Gly	Arg	Leu	Pro	His	Gln	Asn	Leu	His	705	710	715
Leu	Gly	Ser	Val	Lys	Ala	Asn	Val	Gly	His	Gly	Glu	Ser	Ala	Ser	Gly	725	730	735
Ile	Ile	Ala	Leu	Ile	Lys	Thr	Leu	Leu	Met	Met	Glu	Lys	Asn	Met	Ile	740	745	750
Pro	Pro	His	Cys	Gly	Ile	Lys	Thr	Lys	Ile	Asn	His	His	Phe	Pro	Thr	755	760	765
Asp	Leu	Thr	Gln	Arg	Asn	Val	His	Ile	Ala	Lys	Val	Pro	Thr	Ser	Trp	770	775	780
Thr	Arg	Ser	Gly	Gln	Ala	Asn	Pro	Arg	Ile	Ala	Phe	Val	Asn	Asn	Phe	785	790	795
Ser	Ala	Ala	Gly	Gly	Asn	Ser	Ala	Val	Leu	Leu	Gln	Asp	Ala	Pro	Gln	805	810	815
Pro	Ser	Val	Val	Ser	Asp	Val	Thr	Asp	Pro	Arg	Thr	Ser	His	Val	Val	820	825	830
Thr	Met	Ser	Ala	Arg	Ser	Ala	Asp	Ser	Leu	Arg	Lys	Asn	Leu	Ala	Asn	835	840	845
Leu	Lys	Glu	Leu	Val	Glu	Gly	Gln	Gly	Asp	Ser	Glu	Val	Gly	Phe	Leu	850	855	860
Ser	Lys	Leu	Ser	Tyr	Thr	Thr	Thr	Ala	Arg	Arg	Met	His	His	Gln	Phe	865	870	875
Arg	Ala	Ser	Val	Thr	Ala	Gln	Thr	Arg	Glu	Gln	Leu	Leu	Lys	Gly	Leu	885	890	895
Asp	Ser	Ala	Ile	Glu	Arg	Gln	Asp	Val	Lys	Arg	Ile	Pro	Ala	Ala	Ala	900	905	910
Pro	Ser	Val	Gly	Phe	Val	Phe	Ser	Gly	Gln	Gly	Ala	Gln	Tyr	Arg	Gly	915	920	925
Met	Gly	Lys	Glu	Tyr	Phe	Thr	Ser	Phe	Thr	Ala	Phe	Arg	Ser	Glu	Ile	930	935	940
Met	Ser	Tyr	Asp	Ser	Ile	Ala	Gln	Ala	Gln	Gly	Phe	Pro	Ser	Ile	Leu	945	950	955

## 23

Pro Leu Ile Arg Gly Glu Val Glu Ala Asp Ser Leu Ser Pro Val Glu  
 965 970 975  
 Ile Gln Leu Gly Leu Thr Cys Leu Gln Met Ala Leu Ala Lys Leu Trp  
 980 985 990  
 Lys Ser Phe Gly Val Glu Pro Gly Phe Val Leu Gly His Ser Leu Gly  
 995 1000 1005  
 His Tyr Ala Ala Leu His Val Ala Gly Val Leu Ser Ala Asn Asp Thr  
 1010 1015 1020  
 Ile Tyr Leu Thr Gly Ile Arg Ala Gln Leu Leu Val Asp Lys Cys Gln  
 1025 1030 1035 1040  
 Ala Gly Thr His Ser Met Leu Ala Val Arg Ala Ser Leu Leu Gln Ile  
 1045 1050 1055  
 Gln Gln Phe Leu Asp Ala Asn Ile His Glu Val Ala Cys Val Asn Gly  
 1060 1065 1070  
 Ser Arg Glu Val Val Ile Ser Gly Arg Val Ala Asp Ile Asp Gln Leu  
 1075 1080 1085  
 Val Gly Leu Leu Ser Ala Asp Asn Ile Lys Ala Thr Arg Val Lys Val  
 1090 1095 1100  
 Pro Phe Ala Phe His Ser Ala Gln Val Asp Pro Ile Leu Ser Asp Leu  
 1105 1110 1115 1120  
 Asp Thr Ala Ala Ser Arg Val Thr Phe His Ser Leu Gln Ile Pro Val  
 1125 1130 1135  
 Leu Cys Ala Leu Asp Ser Ser Val Ile Ser Pro Gly Asn His Gly Val  
 1140 1145 1150  
 Ile Gly Pro Leu His Leu Gln Arg His Cys Arg Glu Thr Val Asn Phe  
 1155 1160 1165  
 Glu Gly Ala Leu His Ala Ala Glu His Glu Lys Ile Ile Asn Lys Thr  
 1170 1175 1180  
 Ser Thr Leu Trp Ile Glu Ile Gly Pro His Val Val Cys Ser Thr Phe  
 1185 1190 1195 1200  
 Leu Lys Ser Ser Leu Gly Pro Ser Thr Pro Ala Ile Ala Ser Leu Arg  
 1205 1210 1215  
 Arg Asn Asp Asp Cys Trp Lys Val Leu Ala Asp Gly Leu Ser Ser Leu  
 1220 1225 1230  
 Tyr Ser Ser Gly Leu Thr Ile Asp Leu Asn Glu Tyr His Arg Asp Phe  
 1235 1240 1245  
 Lys Ala Ser His Gln Val Leu Arg Leu Pro Cys Tyr Ser Trp Glu His  
 1250 1255 1260

## 24

Lys Asn Tyr Trp Ile Gln Tyr Lys Tyr Asp Trp Ser Leu Ala Lys Gly  
 265 1270 1275 1280  
 Asp Pro Pro Ile Ala Pro Asn Ser Ser Val Glu Ala Val Ser Ala Leu  
 1285 1290 1295  
 Ser Thr Pro Ser Val Gln Lys Ile Leu Gln Glu Thr Ser Leu Asp Gln  
 1300 1305 1310  
 Val Leu Thr Ile Val Ala Glu Thr Asp Leu Ala Ser Pro Leu Leu Ser  
 1315 1320 1325  
 Glu Val Ala Gln Gly His Arg Val Asn Gly Val Lys Val Cys Thr Ser  
 1330 1335 1340  
 Ser Val Tyr Ala Asp Val Gly Leu Thr Leu Gly Lys Tyr Ile Leu Asp  
 345 1350 1355 1360  
 Asn Tyr Arg Thr Asp Leu Glu Gly Tyr Ala Val Asp Val His Gly Ile  
 1365 1370 1375  
 Glu Val His Lys Pro Leu Leu Leu Lys Glu Asp Met Asn Gly Thr Pro  
 1380 1385 1390  
 Gln Ala Thr Pro Phe Arg Ile Glu Val Arg Tyr Pro Ile Gln Ser Thr  
 1395 1400 1405  
 Thr Ala Leu Met Ser Ile Ser Thr Thr Gly Pro Asn Gly Gln His Ile  
 1410 1415 1420  
 Lys His Ala Asn Cys Glu Leu Arg Leu Glu His Pro Ser Gln Trp Glu  
 425 1430 1435 1440  
 Ala Glu Trp Asp Arg Gln Ala Tyr Leu Ile Asn Arg Ser Val Asn Cys  
 1445 1450 1455  
 Leu Leu Gln Arg Ser Ala Gln Gly Leu Asp Ser Met Leu Ala Thr Gly  
 1460 1465 1470  
 Met Val Tyr Lys Val Phe Ser Ser Leu Val Asp Tyr Ala Asp Gly Tyr  
 1475 1480 1485  
 Lys Gly Leu Gln Glu Val Val Leu His Ser Gln Glu Leu Glu Gly Thr  
 1490 1495 1500  
 Ala Lys Val Arg Phe Gln Thr Pro Ser Gly Gly Phe Val Cys Asn Pro  
 505 1510 1515 1520  
 Met Trp Ile Asp Ser Cys Gly Gln Thr Thr Gly Phe Met Met Asn Cys  
 1525 1530 1535  
 His Gln Thr Thr Pro Asn Asp Tyr Val Tyr Val Asn His Gly Trp Lys  
 1540 1545 1550  
 Ser Met Arg Leu Ala Lys Ala Phe Arg Glu Asp Gly Thr Tyr Arg Thr  
 1555 1560 1565

## 25

Tyr Ile Arg Met Arg Pro Ile Asp Ser Thr Lys Phe Ala Gly Asp Leu  
 1570 1575 1580

Tyr Ile Leu Asp Glu Asp Asp Thr Val Val Gly Val Tyr Gly Asp Ile  
 585 1590 1595 1600

Thr Phe Gln Gly Leu Pro Arg Arg Val Leu Asn Thr Val Leu Pro Ser  
 1605 1610 1615

Ala Asn Ala Val Pro Val Asp Ala Pro Met Gly Arg Arg Asp Val Pro  
 1620 1625 1630

Pro Ser Arg Met Asp Val Pro Pro Val Arg Ser Gly Glu Gly Pro Pro  
 . 1635 1640 . 1645

Thr Ser Ala Pro Thr Gln Gln Ala Ile Ala Leu Pro Phe Ala Ala Asp  
 1650 1655 1660

Thr Ser Met Asp Ser Arg Leu Arg Pro Leu Leu Arg Ile Leu Ser Glu  
 665 1670 1675 1680

Glu Ile Gly Leu Gly Leu Asp Val Leu Ser Asp Asp Glu Leu Asp Phe  
 1685 1690 1695

Ala Asp His Gly Val Asp Ser Leu Leu Ser Leu Thr Ile Thr Gly Arg  
 1700 1705 1710

Met Arg Glu Glu Leu Gly Leu Asp Val Glu Ser Thr Ala Phe Met Asn  
 1715 1720 1725

Cys Pro Thr Leu Gly Ser Phe Lys Leu Phe Leu Gly Leu Val Asp Gln  
 1730 1735 1740

Asp Asn Lys Gly Ser Ser Gly Ser Asp Gly Ser Gly Arg Ser Ser Pro  
 745 1750 1755 1760

Ala Pro Gly Thr Glu Ser Gly Ala Thr Thr Pro Pro Met Ser Glu Glu  
 1765 1770 1775

Asp Gln Asp Lys Ile Val Ser Ser His Ser Leu His Gln Phe Gln Ala  
 1780 1785 1790

Ser Ser Thr Leu Leu Gln Gly Ser Pro Ser Lys Ala Arg Ser Thr Leu  
 1795 1800 1805

Phe Leu Leu Pro Asp Gly Ser Gly Ser Ala Thr Ser Tyr Ala Ser Leu  
 1810 1815 1820

Pro Pro Ile Ser Pro Asp Gly Asp Val Ala Val Tyr Gly Leu Asn Cys  
 825 1830 1835 1840

Pro Trp Leu Lys Asp Ser Ser Tyr Leu Val Glu Phe Gly Leu Lys Gly  
 1845 1850 1855

Leu Thr Glu Leu Tyr Val Asn Glu Ile Leu Arg Arg Lys Pro Gln Gly  
 1860 1865 1870

26

Pro Tyr Asn Leu Gly Gly Trp Ser Ala Gly Gly Ile Cys Ala Tyr Glu  
1875 1880 1885

Ala Ala Leu Ile Leu Thr Arg Ala Gly His Gln Val Asp Arg Leu Ile  
1890 1895 1900

Leu Ile Asp Ser Pro Asn Pro Val Gly Leu Glu Lys Leu Pro Pro Arg  
905 1910 1915 1920

Leu Tyr Asp Phe Leu Asn Ser Gln Asn Val Phe Gly Ser Asp Asn Pro  
1925 1930 1935

His Ser Thr Ala Gly Thr Ser Val Lys Ala Pro Glu Trp Leu Leu Ala  
1940 1945 1950

His Phe Leu Ala Phe Ile Asp Ala Leu Asp Ala Tyr Val Ala Val Pro  
1955 1960 1965

Trp Asp Ser Gly Leu Val Gly Leu Ala Ser Pro Leu Pro Ala Pro Pro  
1970 1975 1980

Gln Thr Tyr Met Leu Trp Ala Glu Asp Gly Val Cys Lys Asp Ser Asp  
985 1990 1995 2000

Ser Ala Arg Pro Glu Tyr Arg Asp Asp Asp Pro Arg Glu Met Arg Trp  
2005 2010 2015

Leu Leu Glu Asn Arg Thr Asn Phe Gly Pro Asn Gly Trp Glu Ala Leu  
2020 2025 2030

Leu Gly Gly Lys Glu Gly Leu Phe Met Asp Arg Ile Ala Glu Ala Asn  
2035 2040 2045

His Phe Ser Met Leu Lys Arg Gly Arg Asn Ala Glu Tyr Val Ser Ala  
2050 2055 2060

Phe Leu Ala Arg Ala Leu Asp Asn  
065 2070

<210> 7

<211> 228

<212> DNA

<213> Fusarium graminearum

<220>

<221> CDS

<222> (1) .. (228)

<400> 7

ggc aac aga gcg ttc act cct ggt cga etc aac tac ttc ttc aag ttc 48  
Gly Asn Arg Ala Phe Thr Pro Gly Arg Leu Asn Tyr Phe Phe Lys Phe  
1 5 10 15

agt ggg cct agc gct agt gtt gat acg gct tgc tcc tcc agt etc gtt 96  
Ser Gly Pro Ser Ala Ser Val Asp Thr Ala Cys Ser Ser Ser Leu Val  
20 25 30

atggaggagg	tctacgtctt	cggcgaccaa	acggctgact	gccggggcgtt	cttcaccaag	60
gtctttaccc	ggaaggacaa	tgtcctcttg	cagagcttcc	tggaaagagc	tgggtgaagcc	120
gtcagattcg	aaaaccagaa	ccgatcacat	ccctccaaag	cagtccctaa	tttctctacc	180
atccaagagc	ttgtcgaccg	ttattacaga	ggtgacgcga	aggatgccgc	gaccgagagt	240
gctctcgttt	gcatttcaca	gttttgtcat	ttcattgggtg	ccttttgagga	gagacggcct	300
tcgtatattc	aacctaattc	agatgctcga	ctagtcgggtc	tgtgcacagg	tctaattgct	360
gcgaccgcag	tcgctgcctc	agattccctg	acagccctta	ttcctttggc	ggtcgaagcg	420
gtccggatag	cgttcagagc	tggtgcccac	gtgggaaaag	ttgcacagca	gacagaatgt	480
gactccaaga	ctcaatcgtg	gtccaccatc	gttgtgtccg	acgagaaatc	tgcgcaggag	540
gctctggatg	cctttcacaa	agagaanggt	acatcaccca	tcaaccagtt	gtggatcagt	600
gtttcttccg	ccacttcagt	caccatttagc	gttccgccat	ggaccaaaagc	ccgtcttttn	660
gaagagtcag	agttcttccg	cacacaaaag	agtgtctccag	tgtcaatatt	cgctccntat	720
cacgcacgc	acntacactc	gcaatccgat	ttggacaaaa	tcttacgacc	gcagaccaag	780
acaatctttg	gcaataccac	agtgcggttt	ccgctctgct	caagcgtgac	cggcaaaccg	840

ttcaccgctg	aaaatgggtt	tgagcttctc	caggccgcgt	tgaaagaaat	cataatcgat	900
cccctgcgat	gggacaaggt	gttgaaatat	tgtgcggcag	ggaaggcgctc	ggaagccaag	960
gtattcgctg	ttggaccgac	aaacctggcg	agcagtgtgg	tctccgctct	caaagcctca	1020
acaacgaagg	tcacgctgga	agaccactcc	acctggagta	cagtgccacc	gcagggaacg	1080
cgtcattcaa	aacgagagge	tgatactgct	attgtcggct	tctctggctg	cttccccgac	1140
gccgcagata	atgagctttt	ctggcaattg	ttggaacgtg	gcttggatgt	tcaccggccg	1200
gtccctcccg	acaggttccc	tgtcgagagc	cacactgata	caagtggaaa	gaagaagaac	1260
acctctcaca	caccgttttg	caattttatt	gagaagccag	gcttgtttga	cgcgagggttc	1320
ttcaatatgt	caccaagaga	agccgcacaa	actgacccca	tgacgaggct	gatgttgacc	1380
accggctatg	angctatgga	aatggctggc	atcgccctcg	gcanaacacc	ctcnacnaan	1440
catgacagaa	ttggcacttt	ctacggtcaa	acctcanacn	attggcgatga	ggtcaatgcc	1500
gcccangata	ttgatacata	tttcatttca	gggtggcggtg	gcgccttcgg	acctggctcg	1560
atcaactact	tttttaagtt	ttccggacct	antttctcgg	tggacatgcn	tgctaatacca	1620
gcttggccgc	ntatgaatgt	cgcaatcaca	tctctcagag	caaatagagt	cgatactgtg	1680
ttcacgggtg	gtgccaatgt	gtcacgaac	tctgacattt	tctctggctct	cagcagaggc	1740
cacttcttgt	ccaaaaccgg	gtcttgcaag	acatgggata	atgacgctga	tggctactgc	1800
cgcgggcgatg	gtgtttgcac	gggtcatcatg	aagcgtcttg	acgatgccct	tgcggatcga	1860
gaccctgttc	tgggtgttat	ccgcggtatt	ggcacgaacc	attcagcaga	agccgtttct	1920
atcacccatc	cctgtgcgga	gaatcaggca	ttctgtgtcg	acaagggtgt	caaagagtgc	1980
aacgttccact	gcaatgacgt	caactacgtt	gagatgcacg	gcacaggcac	gcaggcaggc	2040
gacggcatcg	agatggaatc	tgtctcatct	gtttttgccc	cacgacagcc	togaagacgt	2100
cctgaccagc	ctctctacgt	tgggtgcagtc	aagtcaaaca	tcggtcacgg	tgaggccgtt	2160
tctggcgat	ccgctctgat	caagggtgtc	cttatgctgc	aaaagaacaa	gatcccgccct	2220
cataccggca	tcaaaaaaca	gatcaacaag	aactttgcac	cagatctcaa	ggagcgcaac	2280
gtcaatatgt	ccttccagac	aacgcctttt	ccacgcccac	ctggtggcaa	gcgaactgtg	2340
ttcatcaaca	acttttagtgc	agctgggtgg	aacacagcga	tgcttttgca	agatggcccc	2400
gagggtcccaa	cggagcccttc	ttctgaccca	agatcgaccc	atgtggtgac	antctcagcg	2460
aaatcttttg	ccgcttttaa	gaggaccctt	gccaaatacg	aagcntattt	gaatgctcac	2520
ccgaatgtcg	gccttcccga	ccttgcatat	actgtcacgg	cccgcggggc	gcattacagc	2580
ttaccgcgcc	gcttcccagt	tcagtcaatt	tctcagttgc	aagcttcgtt	gagggccatt	2640
caggaccaga	cccacaaccc	gateccccct	gcctcacctc	agattgctat	ggctttttaca	2700
gggcaagggt	cacagtacac	tgggatgggc	cagaaattgt	ttgagacatc	gaagcaattc	2760
cgacaagaca	togaagagtt	caatgagatc	gctcttcggc	aggggctccc	ttctattatg	2820
cccttgatcg	atggatctgt	ggaagttcag	catctacgcg	ccacagtcgt	ccaactcggg	2880
atgtgctgta	ttcaaattggc	tcttactcat	ctttgggtcga	cttgggggaat	tcagcccagt	2940
gttgttatcg	gtcactccct	gggagaatac	gctgctctcc	aagccgctgg	cgttctgttc	3000
attgccgaca	ccatctatct	ggttggtaag	cgagctcagc	tgctggagca	gaagtgcaca	3060
gcaggcactc	atgcgatgct	tgccgtgaga	tctcctgttg	gtgggctcca	ggatgtgggt	3120
gccaatagcc	atggcaagat	tgaataattgc	ggcatcaacg	gcgtgtctga	cactgttctt	3180
tccggcacia	tgggcgatat	tgataccgtc	gctcaaaaac	tggcggacgc	tggccagaag	3240
tgcaccaagt	tgaagttgcc	attcgcattt	cattcatccc	aagtcgatcc	cattttggca	3300
gacttcgaaa	aattggcctc	ttcgggtcaat	taccacccac	cccgtgttcc	tgtcatctcg	3360
cccctgttga	gcgatgttgt	cagtgtcggg	ggcgtcttcg	atgcattcta	cctcagtaga	3420
cactgccgca	agacagtcga	ctttgtcggg	gggttgtccg	ctggatgttc	aaccgccacc	3480
attagcgata	cttcgctttg	gctcgagggt	ggcgggtcac	ccctctgcgc	cagcatgatc	3540
aaatcttgct	tgtcagtcct	aacctgggcc	acgatgcggc	gtgatgaaga	cccctggaag	3600
atcattttctg	cgtccatggc	cggcctctat	acggccggga	agtctttgaa	ctgggatgca	3660
ttccacaagg	agaacgagtc	cttgcgcgtg	ctcaacgact	tgccattcta	tggattcgac	3720
gagaagaatt	actgggttaca	gtacactgga	gattggcttt	tgtacaaggg	cgactacccc	3780

aaggcgattg	ccccagctcc	agccgcccgt	gcagcagcaa	ggcctgccaa	agcgaggaaa	3840
tatctgtcaa	cctccgttca	aggtattgtc	tccgaagagg	tcaagggcaa	gactgtcacg	3900
attgtcgccg	agtcgcactt	tgcgcacccg	aaattgttcc	cggtcattgc	cggtcacctt	3960
gtcaacggca	gtgggtctctg	cccgtccaca	ttgtatgccg	acatggccta	cacccttggg	4020
caactaggcg	ttggcctcct	gaagccgggt	gagaagggtg	acatcaatat	tggcacaatg	4080
gacaatccgg	cacctcttct	gctcaagaac	atcaatcagc	ctgaaagcca	aattgttcaa	4140
atgacaatga	agatcgatct	cgatgccagg	aaagcggact	tcgctgtgac	cagcaacaat	4200
ggaaagaaag	atgtcactca	tgccaagtgc	gtgatcgctc	tcgaagatgc	tgctgtttgg	4260
aaagaacaat	ggagcaagac	ttcttacctg	atccagtcga	gaatcgacat	gttgaaacac	4320
aagatggaga	atgggtgaagc	cgacaagggt	agtcgcgcta	tggcttacaa	gctcttcggc	4380
gccctcgttg	actacgccga	tatattccaa	ggcatgcaaa	acgtgggtctt	tgatggaccg	4440
gaattcgagg	cgacgagcaa	catcaaattc	cgcgctgggc	cgaacgatgg	cgacttttac	4500
ttcagcccc	atttcattga	cagcgcttgc	cacctctcan	gcttcaccgt	cnatgctacc	4560
gtcnaccac	aagatgagtg	ctacatttct	catgggtggg	ccagtctccg	cttcattgag	4620
ccgcttcagc	atgaccagca	atactacgca	tacctcaaga	tgcagcccgt	ttcaggaagc	4680
aagatnagag	ctgggtgatgt	ttatgtcttc	aacgcggata	aagangtctt	ccgtcttgcg	4740
gggtggtgtc	gattccagtg	catcccccgc	aagctcatgg	atgtgatgat	gccccaacct	4800
aaggccaatg	caaaggctgc	ttccccagca	tctgcgacc	cgaagcagc	tccaactgtc	4860
aaagcctcca	gcaacctcgc	tgctcctgct	tccgtggaac	tgcccactcc	caagggtcaag	4920
gcgaagaagc	cgggttaaggc	acaaaagatc	aaggcaccag	caccacccaa	gcaagcctct	4980
tctggtagtc	ttgtcgtccg	ggctttcgaa	atcatcgcaa	aggagattga	tgctcgaccaa	5040
tctgaactca	atgacgatat	ccaatgggct	gacatgggtg	ttgacagctt	gatgtcgctg	5100
accatctctg	gcaagttcag	ggaagatctt	gacctcgagg	ttgngagcac	cctcttcacc	5160
gactacgcct	ccgtggggcgc	ccttcgggaag	catctcagcg	gcatgtctgg	tcctgagccc	5220
attgctactg	gtgatgcaag	ctcgggtgaa	tctacggact	cgggttcgga	gagcgacgag	5280
gataccgtcg	agtcgggcat	caccaccccg	gacaccgaag	antccccagc	gaagcctcag	5340
gaacagggga	aatccgctgc	tgtggaggcg	atgggtcaac	caccttcgc	cgaagggcag	5400
gatatgattg	agaccattcg	ngtagtcac	gccaagaaa	tggagatgga	tctggctgaa	5460
atcaccggacg	ntactgacct	cagtaacctc	ggaatggact	ccttgatggc	cctgaccgtt	5520
cttggggaagc	ttcgagagga	tcatgacatt	gatttggatc	cgaccatctt	ggcggacaat	5580
ccaaccctgg	cgcacctccg	caaagcattg	ggtcttgaga	aagccaagcc	ggcgcctgca	5640
ccgaacaacg	angtcaggac	caatgtcggt	gtcgtcctcg	ctgcgcctcc	ggttgangtt	5700
gtcgtgcana	tnccccccgc	cacatctgtg	ttgcttcagg	gcaatcccaa	gacagccacc	5760
aanaatctgt	tcttggtccc	tgatggctcc	gggtctgcca	catcctatgt	gtctatccc	5820
gccattgaca	gcaanaacct	ggctgtctat	gggttaaatt	gtccattcat	gaaagatccc	5880
accagctata	cttgccgggat	cnaaagtgtc	tccaanctgt	atcttgaaaa	agtccttcnc	5940
cgacagccta	atgggtccata	tatcctcngg	ggctgggtcag	caagtgggtg	ctttgcttac	6000
naaatcacca	nacaactcan	tgacttacag	aanctccacc	ctgacaagaa	ctacaccgtt	6060
gaaaaattga	atcttatena	ctcnccatgc	ccgatccgtc	tggagcctct	acctgcccgc	6120
ctgcaccact	tctttgatga	aattgggtctt	ttggggacag	gtaccggcaa	gaccccgaa	6180
tggctgctac	cccatttcga	atacagcatc	aaggcgttga	ctgcctaccg	ccctgagctg	6240
aagtctacac	gtgatttcaa	tgcaccaccc	accttgcctc	tctggggccac	tgatgggtgtc	6300
tgccggcaagc	ccggtgaccc	gagaccgcgc	ccacaagcgg	atgatccgaa	gagcatgaag	6360
tggctgctgg	aaaaccgcac	ggactttgga	cccaatggat	gggacaagct	tttgggtgct	6420
gaagtctgta	aaatgggttac	agtgggtggg	aaccatttca	ccatgatgaa	accaccgggtg	6480
gctaagggtg	ttgggtcaata	catccgggag	tctctctcca	tgtanagagc	ttga	6534

30

&lt;212&gt; PRT

&lt;213&gt; Wagiella dermatidis

&lt;400&gt; 10

Met Glu Glu Val Tyr Val Phe Gly Asp Gln Thr Ala Asp Cys Arg Ala  
1 5 10 15

Phe Phe Thr Lys Val Phe Thr Arg Lys Asp Asn Val Leu Leu Gln Ser  
20 25 30

Phe Leu Glu Arg Ala Gly Glu Ala Val Arg Phe Glu Asn Gln Asn Arg  
35 40 45

Ser His Pro Ser Lys Ala Val Pro Asn Phe Ser Thr Ile Gln Glu Leu  
50 55 60

Val Asp Arg Tyr Tyr Arg Gly Asp Ala Lys Asp Ala Ala Thr Glu Ser  
65 70 75 80

Ala Leu Val Cys Ile Ser Gln Phe Cys His Phe Ile Gly Ala Phe Glu  
85 90 95

Glu Arg Arg Pro Ser Tyr Ile Gln Pro Asn Ser Asp Ala Arg Leu Val  
100 105 110

Gly Leu Cys Thr Gly Leu Ile Ala Ala Thr Ala Val Ala Ala Ser Asp  
115 120 125

Ser Leu Thr Ala Leu Ile Pro Leu Ala Val Glu Ala Val Arg Ile Ala  
130 135 140

Phe Arg Ala Gly Ala His Val Gly Lys Val Ala Gln Gln Thr Glu Cys  
145 150 155 160

Asp Ser Lys Thr Gln Ser Trp Ser Thr Ile Val Ala Ala Asp Glu Lys  
165 170 175

Ser Ala Gln Glu Ala Leu Asp Ala Phe His Lys Glu Xaa Gly Thr Ser  
180 185 190

Pro Ile Asn Gln Leu Trp Ile Ser Val Ser Ser Ala Thr Ser Val Thr  
195 200 205

Ile Ser Val Pro Pro Trp Thr Lys Ala Arg Leu Xaa Glu Glu Ser Glu  
210 215 220

Phe Phe Arg Thr Gln Lys Ser Ala Pro Val Ser Ile Phe Ala Pro Tyr  
225 230 235 240

His Ala Ser His Xaa His Ser Gln Ser Asp Leu Asp Lys Ile Leu Arg  
245 250 255

Pro Gln Thr Lys Thr Ile Phe Gly Asn Thr Thr Val Arg Phe Pro Val  
260 265 270

## 31

Cys Ser Ser Val Thr Gly Lys Pro Phe Asn Ala Glu Asn Gly Phe Glu	275	280	285
Leu Leu Gln Ala Ala Leu Lys Glu Ile Ile Ile Asp Pro Leu Arg Trp	290	295	300
Asp Lys Val Leu Lys Tyr Cys Ala Ala Gly Lys Ala Ser Glu Ala Lys	305	310	315 320
Val Phe Ala Val Gly Pro Thr Asn Leu Ala Ser Ser Val Val Ser Ala	325	330	335
Leu Lys Ala Ser Thr Thr Lys Val Thr Leu Glu Asp His Ser Thr Trp	340	345	350
Ser Thr Val Pro Pro Gln Gly Thr Arg His Ser Lys Arg Glu Ala Asp	355	360	365
Ile Ala Ile Val Gly Phe Ser Gly Arg Phe Pro Asp Ala Ala Asp Asn	370	375	380
Glu Leu Phe Trp Gln Leu Leu Glu Arg Gly Leu Asp Val His Arg Pro	385	390	395 400
Val Pro Pro Asp Arg Phe Pro Val Glu Ser His Thr Asp Pro Ser Gly	405	410	415
Lys Lys Lys Asn Thr Ser His Thr Pro Phe Gly Asn Phe Ile Glu Lys	420	425	430
Pro Gly Leu Phe Asp Ala Arg Phe Phe Asn Met Ser Pro Arg Glu Ala	435	440	445
Ala Gln Thr Asp Pro Met Gln Arg Leu Met Leu Thr Thr Gly Tyr Xaa	450	455	460
Ala Met Glu Met Ala Gly Ile Val Pro Gly Xaa Thr Pro Ser Thr Xaa	465	470	475 480
His Asp Arg Ile Gly Thr Phe Tyr Gly Gln Thr Ser Xaa Xaa Trp Arg	485	490	495
Glu Val Asn Ala Ala Xaa Asp Ile Asp Thr Tyr Phe Ile Ser Gly Gly	500	505	510
Val Arg Ala Phe Gly Pro Gly Xaa Ile Asn Tyr Phe Phe Lys Phe Ser	515	520	525
Gly Pro Xaa Phe Ser Val Asp Met Xaa Ala Asn Pro Ala Trp Pro Xaa	530	535	540
Met Asn Val Ala Ile Thr Ser Leu Arg Ala Asn Glu Cys Asp Thr Val	545	550	555 560
Phe Thr Gly Gly Ala Asn Val Leu Thr Asn Ser Asp Ile Phe Ser Gly	565	570	575

## 32

Leu	Ser	Arg	Gly	His	Phe	Leu	Ser	Lys	Thr	Gly	Ser	Cys	Lys	Thr	Trp
			580					585					590		
Asp	Asn	Asp	Ala	Asp	Gly	Tyr	Cys	Arg	Gly	Asp	Gly	Val	Cys	Thr	Val
			595				600					605			
Ile	Met	Lys	Arg	Leu	Asp	Asp	Ala	Leu	Ala	Asp	Arg	Asp	Pro	Val	Leu
	610					615					620				
Gly	Val	Ile	Arg	Gly	Ile	Gly	Thr	Asn	His	Ser	Ala	Glu	Ala	Val	Ser
625					630					635					640
Ile	Thr	His	Pro	Cys	Ala	Glu	Asn	Gln	Ala	Phe	Leu	Phe	Asp	Lys	Val
				645					650					655	
Leu	Lys	Glu	Cys	Asn	Val	His	Cys	Asn	Asp	Val	Asn	Tyr	Val	Glu	Met
			660					665					670		
His	Gly	Thr	Gly	Thr	Gln	Ala	Gly	Asp	Gly	Ile	Glu	Met	Glu	Ser	Val
		675					680					685			
Ser	Ser	Val	Phe	Ala	Pro	Arg	Gln	Pro	Arg	Arg	Arg	Pro	Asp	Gln	Pro
	690					695					700				
Leu	Tyr	Val	Gly	Ala	Val	Lys	Ser	Asn	Ile	Gly	His	Gly	Glu	Ala	Val
705					710					715					720
Ser	Gly	Val	Ser	Ala	Leu	Ile	Lys	Val	Leu	Leu	Met	Leu	Gln	Lys	Asn
			725						730					735	
Lys	Ile	Pro	Pro	His	Thr	Gly	Ile	Lys	Lys	Gln	Ile	Asn	Lys	Asn	Phe
			740					745					750		
Ala	Pro	Asp	Leu	Lys	Glu	Arg	Asn	Val	Asn	Ile	Ala	Phe	Gln	Thr	Thr
		755					760					765			
Pro	Phe	Pro	Arg	Pro	Pro	Gly	Gly	Lys	Arg	Thr	Val	Phe	Ile	Asn	Asn
	770					775					780				
Phe	Ser	Ala	Ala	Gly	Gly	Asn	Thr	Ala	Met	Leu	Leu	Gln	Asp	Gly	Pro
785					790					795					800
Glu	Val	Pro	Thr	Glu	Pro	Ser	Ser	Asp	Pro	Arg	Ser	Thr	His	Val	Val
				805					810					815	
Thr	Xaa	Ser	Ala	Lys	Ser	Leu	Ala	Ala	Phe	Lys	Arg	Thr	Leu	Ala	Lys
			820					825					830		
Tyr	Glu	Ala	Tyr	Leu	Asn	Ala	His	Pro	Asn	Val	Gly	Leu	Pro	Asp	Leu
	835						840					845			
Ala	Tyr	Thr	Val	Thr	Ala	Arg	Arg	Ala	His	Tyr	Ser	Leu	Pro	Arg	Arg
	850					855					860				
Phe	Pro	Val	Gln	Ser	Ile	Ser	Gln	Leu	Gln	Ala	Ser	Leu	Arg	Ala	Ile
865					870					875					880

33

Gln Asp Gln Thr His Asn Pro Ile Pro Leu Ala Ser Pro Gln Ile Ala  
885 890 895

Met Ala Phe Thr Gly Gln Gly Ser Gln Tyr Thr Gly Met Gly Gln Lys  
900 905 910

Leu Phe Glu Thr Ser Lys Gln Phe Arg Gln Asp Ile Glu Glu Phe Asn  
915 920 925

Glu Ile Ala Leu Arg Gln Gly Leu Pro Ser Ile Met Pro Leu Ile Asp  
930 935 940

Gly Ser Val Glu Val Gln His Leu Pro Pro Thr Val Val Gln Leu Gly  
945 950 955 960

Met Cys Cys Ile Gln Met Ala Leu Thr His Leu Trp Ser Thr Trp Gly  
965 970 975

Ile Gln Pro Ser Val Val Ile Gly His Ser Leu Gly Glu Tyr Ala Ala  
980 985 990

Leu Gln Ala Ala Gly Val Leu Ser Ile Ala Asp Thr Ile Tyr Leu Val  
995 1000 1005

Gly Lys Arg Ala Gln Leu Leu Glu Gln Lys Cys Thr Ala Gly Thr His  
1010 1015 1020

Ala Met Leu Ala Val Arg Ser Pro Val Gly Gly Leu Gln Asp Val Val  
1025 1030 1035 1040

Ala Asn Ser His Gly Lys Ile Glu Asn Cys Gly Ile Asn Gly Val Ser  
1045 1050 1055

Asp Thr Val Leu Ser Gly Thr Met Gly Asp Ile Asp Thr Val Ala Gln  
1060 1065 1070

Lys Leu Ala Asp Ala Gly Gln Lys Cys Thr Lys Leu Lys Leu Pro Phe  
1075 1080 1085

Ala Phe His Ser Ser Gln Val Asp Pro Ile Leu Ala Asp Phe Glu Lys  
1090 1095 1100

Leu Ala Ser Ser Val Asn Tyr His Pro Pro Arg Val Pro Val Ile Ser  
1105 1110 1115 1120

Pro Leu Leu Ser Asp Val Val Ser Val Gly Gly Val Phe Asp Ala Phe  
1125 1130 1135

Tyr Leu Ser Arg His Cys Arg Lys Thr Val Asp Phe Val Gly Gly Leu  
1140 1145 1150

Ser Ala Gly Met Ser Thr Ala Thr Ile Ser Asp Thr Ser Leu Trp Leu  
1155 1160 1165

Glu Val Gly Gly His Pro Leu Cys Ala Ser Met Ile Lys Ser Cys Leu  
1170 1175 1180

## 34

Ser Val Pro Thr Leu Ala Thr Met Arg Arg Asp Glu Asp Pro Trp Lys  
1185 1190 1195 1200  
Ile Ile Ser Ala Ser Met Ala Gly Leu Tyr Thr Ala Gly Lys Ser Leu  
1205 1210 1215  
Asn Trp Asp Ala Phe His Lys Glu Asn Glu Ser Leu Arg Val Leu Asn  
1220 1225 1230  
Asp Leu Pro Phe Tyr Gly Phe Asp Glu Lys Asn Tyr Trp Leu Gln Tyr  
1235 1240 1245  
Thr Gly Asp Trp Leu Leu Tyr Lys Gly Asp Tyr Pro Lys Ala Ile Ala  
1250 1255 1260  
Pro Ala Pro Ala Ala Ala Ala Ala Arg Pro Ala Lys Ala Arg Lys  
1265 1270 1275 1280  
Tyr Leu Ser Thr Ser Val Gln Gly Ile Val Ser Glu Glu Val Lys Gly  
1285 1290 1295  
Lys Thr Val Thr Ile Val Ala Glu Ser Asp Phe Ala His Pro Lys Leu  
1300 1305 1310  
Phe Pro Val Ile Ala Gly His Leu Val Asn Gly Ser Gly Leu Cys Pro  
1315 1320 1325  
Ser Thr Leu Tyr Ala Asp Met Ala Tyr Thr Leu Gly Gln Leu Gly Val  
1330 1335 1340  
Gly Leu Leu Lys Pro Gly Glu Lys Val Asp Ile Asn Ile Gly Thr Met  
1345 1350 1355 1360  
Asp Asn Pro Ala Pro Leu Leu Leu Lys Asn Ile Asn Gln Pro Glu Ser  
1365 1370 1375  
Gln Ile Val Gln Met Thr Met Lys Ile Asp Leu Asp Ala Arg Lys Ala  
1380 1385 1390  
Asp Phe Ala Val Thr Ser Asn Asn Gly Lys Lys Asp Val Thr His Ala  
1395 1400 1405  
Lys Cys Val Ile Val Phe Glu Asp Ala Ala Val Trp Lys Glu Gln Trp  
1410 1415 1420  
Ser Lys Thr Ser Tyr Leu Ile Gln Ser Arg Ile Asp Met Leu Lys His  
1425 1430 1435 1440  
Lys Met Glu Asn Gly Glu Ala Asp Lys Val Ser Arg Ala Met Ala Tyr  
1445 1450 1455  
Lys Leu Phe Gly Ala Leu Val Asp Tyr Ala Asp Ile Phe Gln Gly Met  
1460 1465 1470  
Gln Asn Val Val Phe Asp Gly Pro Glu Phe Glu Ala Thr Ser Asn Ile  
1475 1480 1485

## 35

Lys Phe Arg Ala Gly Pro Asn Asp Gly Asp Phe Tyr Phe Ser Pro Tyr  
 1490 1495 1500  
 Phe Ile Asp Ser Ala Cys His Leu Ser Xaa Phe Thr Val Xaa Ala Thr  
 1505 1510 1515 1520  
 Val Xaa Pro Gln Asp Glu Cys Tyr Ile Ser His Gly Trp Ser Ser Leu  
 1525 1530 1535  
 Arg Phe Ile Glu Pro Leu Gln His Asp Gln Gln Tyr Tyr Ala Tyr Leu  
 1540 1545 1550  
 Lys Met Gln Pro Val Ser Gly Ser Lys Xaa Arg Ala Gly Asp Val Tyr  
 1555 1560 1565  
 Val Phe Asn Ala Asp Lys Xaa Val Leu Arg Leu Ala Gly Gly Val Arg  
 1570 1575 1580  
 Phe Gln Cys Ile Pro Arg Lys Leu Met Asp Val Met Met Pro Lys Pro  
 1585 1590 1595 1600  
 Lys Ala Asn Ala Lys Ala Ala Ser Pro Ala Ser Ala Thr Pro Lys Ala  
 1605 1610 1615  
 Ala Pro Thr Val Lys Ala Ser Ser Asn Leu Ala Ala Pro Ala Ser Val  
 1620 1625 1630  
 Glu Leu Pro Thr Pro Lys Val Lys Ala Lys Lys Pro Val Lys Ala Gln  
 1635 1640 1645  
 Lys Ile Lys Ala Pro Ala Pro Pro Lys Gln Ala Ser Ser Gly Ser Leu  
 1650 1655 1660  
 Val Val Arg Ala Phe Glu Ile Ile Ala Lys Glu Ile Asp Val Asp Gln  
 1665 1670 1675 1680  
 Ser Glu Leu Asn Asp Asp Ile Gln Trp Ala Asp Met Gly Val Asp Ser  
 1685 1690 1695  
 Leu Met Ser Leu Thr Ile Ser Gly Lys Phe Arg Glu Asp Leu Asp Leu  
 1700 1705 1710  
 Glu Val Xaa Ser Thr Leu Phe Thr Asp Tyr Ala Ser Val Gly Ala Leu  
 1715 1720 1725  
 Arg Lys His Leu Ser Gly Met Ser Gly Pro Glu Pro Ile Ala Thr Gly  
 1730 1735 1740  
 Asp Ala Ser Ser Val Glu Ser Thr Asp Ser Gly Ser Glu Ser Asp Glu  
 1745 1750 1755 1760  
 Asp Thr Val Glu Ser Gly Ile Thr Thr Pro Asp Thr Glu Xaa Phe Pro  
 1765 1770 1775  
 Ala Lys Xaa Gln Glu Gln Gly Lys Ser Ala Ala Val Glu Ala Met Ala  
 1780 1785 1790

36

Gln Pro Pro Ser Ala Glu Gly Gln Asp Met Ile Glu Thr Ile Arg Val  
1795 1800 1805

Val Ile Ala Gln Glu Met Glu Met Asp Leu Ala Glu Ile Thr Asp Xaa  
1810 1815 1820

Thr Asp Leu Ser Asn Leu Gly Met Asp Ser Leu Met Ala Leu Thr Val  
1825 1830 1835 1840

Leu Gly Lys Leu Arg Glu Asp His Asp Ile Asp Leu Asp Pro Thr Ile  
1845 1850 1855

Leu Ala Asp Asn Pro Thr Leu Ala His Leu Arg Lys Ala Leu Gly Leu  
. 1860 . 1865 . 1870

Glu Lys Ala Lys Pro Ala Pro Ala Pro Lys Gln Xaa Val Arg Thr Asn  
1875 1880 1885

Val Val Val Ala Pro Ala Ala Pro Pro Val Xaa Val Val Val Xaa Xaa  
1890 1895 1900

Pro Pro Ala Thr Ser Val Leu Leu Gln Gly Asn Pro Lys Thr Ala Thr  
1905 1910 1915 1920

Xaa Asn Leu Phe Leu Phe Pro Asp Gly Ser Gly Ser Ala Thr Ser Tyr  
1925 1930 1935

Val Ser Ile Pro Ala Ile Asp Ser Xaa Asn Leu Ala Val Tyr Gly Leu  
1940 1945 1950

Asn Cys Pro Phe Met Lys Asp Pro Thr Ser Tyr Thr Cys Gly Ile Xaa  
1955 1960 1965

Ser Val Ser Xaa Leu Tyr Leu Glu Lys Val Leu Xaa Arg Gln Pro Asn  
1970 1975 1980

Gly Pro Tyr Ile Leu Xaa Gly Trp Ser Ala Ser Gly Val Phe Ala Tyr  
1985 1990 1995 2000

Xaa Ile Thr Xaa Gln Leu Xaa Asp Leu Gln Xaa Leu His Pro Asp Lys  
2005 2010 2015

Asn Tyr Thr Val Glu Lys Leu Asn Leu Ile Xaa Ser Pro Cys Pro Ile  
2020 2025 2030

Arg Leu Glu Pro Leu Pro Ala Arg Leu His His Phe Phe Asp Glu Ile  
2035 2040 2045

Gly Leu Leu Gly Thr Gly Thr Gly Lys Thr Pro Asn Trp Leu Leu Pro  
2050 2055 2060

His Phe Glu Tyr Ser Ile Lys Ala Leu Thr Ala Tyr Arg Pro Glu Leu  
2065 2070 2075 2080

Lys Ser Thr Arg Asp Phe Asn Ala Pro Pro Thr Leu Leu Ile Trp Ala  
2085 2090 2095

37

Thr Asp Gly Val Cys Gly Lys Pro Gly Asp Pro Arg Pro Pro Pro Gln  
2100 2105 2110

Ala Asp Asp Pro Lys Ser Met Lys Trp Leu Leu Glu Asn Arg Thr Asp  
2115 2120 2125

Phe Gly Pro Asn Gly Trp Asp Lys Leu Leu Gly Ala Glu Val Cys Lys  
2130 2135 2140

Met Val Thr Val Val Gly Asn His Phe Thr Met Met Lys Pro Pro Val  
2145 2150 2155 2160

Ala Lys Gly Val Gly Gln Tyr Ile Arg Glu Ser Leu Ser Met Xaa Arg  
2165 2170 2175

Ala

<210> 11

<211> 6330

<212> DNA

<213> Aspergillus parasiticus

<400> 11

```

atggctcaat caaggcaact cttctcttc ggcatcaga cagcggattt tgttcccaag 60
ctcgcagtt tactatccgt ccaggacagc cctattctag ccgcctttct tgaccagtcc 120
cactatgtcg tgcgagccca gatgctgcag agcatgaaca cggttgatca caagttggct 180
cgaaccgctg acctgcgcca aatggtccag aagtatgtcg acggcaaact gacccttgca 240
tttcgaaccg ctctagtgtg cctctgccag ttgggatgct tcatccggga atatgaggaa 300
tctggcaaca tgtaccaca gccagtgac agctacgtgc tgggattctg catgggttcc 360
ttggccgctg tggcggttaag ctgcagtcgc tccctgtcag agctgctgcc tatcgtgtta 420
caaactgtgt tgattgcctt ccgcctcggg ctttgcgcc tggagatgcg ggatcgggtg 480
gatgggtgta gcgatgatcg aggtgacct ttggtctacca ttgtttgggg tctggatccc 540
cagcaagctc gtgatcagat tgaagtgttc tgcggacca caaacgtacc tcagacaagg 600
cgtccgtgga tcagctgcat ctctaagaat gccatcacc taagtggcag tccatccact 660
ttgagggcgt tctgtgcgat gcctcagatg gccagcacc ggactgccc aattcccatc 720
tgtttaccgg ccacaatgg cgccctcttc acgcaggcag atatcactac catactagac 780
acgacgccta ccactccttg ggagcaactg cccggccaaa taccttata tcccatgtc 840
acggggaatg tagtccagac ttccaactac cgggacctta tagaggtagc cctgtctgag 900
actctcttgg agcaagtgcg actagacttg gttgagactg gactgccacg ccttttgcaa 960
tctcgtcagg tcaagagcgt caccatcgta ccattcttga ctgcgatgaa tgagacaatg 1020
agcaacattc tcccagacag ctttatcagt acagagacaa ggactgacac cggacgagcc 1080
atcccagctt caggtcgacc aggcgcaggc aagtgcaggc tggctattgt gtccatgtcg 1140
gggaggttcc ctgaatcacc gaccaccgaa agcttttggg acctctata caaagggttg 1200
gatgtttgta aagaggttcc ccgtcgacgg tgggacatca acacgcattg ggatcccagc 1260
gggaaagcac gaaacaaagg ggctacaaaa tggggctgct ggetagattt ctcaggcgat 1320
tttgatcccc gattcttctg gatctcgccc aaagaggcgc cacagatgga tccagctcag 1380
cgcatggcct tgatgtctac ttacgaggca atggagcggg ctggtttggg tcccagacac 1440
acgccgtcga occagcgaga ccgcattggg gtcttccacg gagtcaccag taacgactgg 1500
atggagacca atacagccca gaacattgac acatacttca tcaccggtgg aaatcggggg 1560
tttattcccg ggcgcattaa cttctgtttc gaatttgccg gaccagcta taccaatgac 1620
acggcctgtt catccagctc agctgccatc caccctggcct gcaattctct ctggcggggc 1680

```

gactgtgaca	cgggcgtggc	aggaggaact	aacatgatct	atactcctga	tggtcacaca	1740
ggattggaca	aagggttctt	tctttcccgg	actggcaact	gcaaacccta	cgacgacaag	1800
gccgatgggt	actgcccagc	tgaggggggtc	gggacgggtg	tcatacaaac	gctggaagat	1860
gctctggcag	ataatgaccc	caccccttggc	gttattctag	atgctaaaac	taatcactca	1920
gccatgtcgg	agtcgatgac	tcggccgcac	gtggggcgcc	aaatcgataa	catgacggcg	1980
gcgctgaata	ccactggact	ccatcccaat	gacttttagct	acattgagat	gcatggcact	2040
ggcaccaggg	taggggatgc	ggtggagatg	gagtcagtc	tgctgggtgt	tgccgcgtcc	2100
gaaaccgcca	gaaaggcgga	tcagccacta	tttgcggcct	cagccaaggc	caacgtagga	2160
catggagagg	gagtgtctgg	ggttacgagc	cttattaagg	ttctgatgat	gatgcagcac	2220
gataccatac	ctcctcactg	cggcatcaaa	ccgggcagca	aaatcaaccg	caacttccct	2280
gatcttggag	ctcgcaatgt	gcacatcgcc	tttgaaccca	agccctggcc	acgaacacac	2340
actccgcgca	gggtgcttat	caacaacttc	agtgcgcggg	gagggaatac	tgcccttgata	2400
gtggaagacg	ctccggagcg	tcactggccg	acagagaagg	atccgcgctc	tagtcataac	2460
gtcgccctgt	ctgcgcgatg	gggggcttcc	atgaaaacca	acctcgaaac	actgcatcag	2520
tatctcctga	aaaaccccca	cactgatctc	gcgcagctgt	catatactac	tactgcgcgt	2580
cgaaggcatt	atctacaccg	agtgagcgct	actggcgctg	ctgttgaaag	agtgactcgc	2640
aagctagaga	tgcccatata	gaacggggac	ggagtcagtc	gacccaaaag	caagccgaag	2700
attctctttg	ctttcaccgg	acaagggtct	caatatgcaa	ctatgggtta	gcaggtgtac	2760
gatgcgtatc	catctttcag	agaggacctg	gagaagtttg	atcggttggc	gcaaagtcac	2820
ggcttcccta	gctttcttca	cgtctgtact	tcacctaaag	gggatgtgga	agagatggct	2880
cccgttgttg	tgcaactggc	tatcacttgt	ctccaaatgg	cccttactaa	cctcatgacc	2940
tccttcggga	tccttcocga	tgtgacagtg	gggcatagtt	tggttgaaat	tgacgccctg	3000
tatggggcgg	gagttctttc	ggcctcagac	gtcgtttacc	ttgttggtca	aagagcggag	3060
ctactccagg	agcgtgcca	acgcgggacg	catgccatgc	tggtgtgtaa	agctaccctt	3120
gaagcgttgt	cccaatggat	ccaggatcat	gactgtgagg	tgccctgtat	taatggccct	3180
gaagataccg	ttctcagtg	caccactaag	aatgttgccg	aggttcaacg	cgtatgacg	3240
gacaacggga	tcaaatgcac	gctgltgaaa	ctgccgtttg	ccttccattc	tgcccagggtg	3300
caacotattc	tgagcgaact	tgaggccctg	gctcaggagg	cgacatttgc	caagcctcaa	3360
ctactaattc	tctctccctt	gctgcggaca	gaaatccacg	aacaaggcgt	cgtgactcca	3420
tcatatgtcg	cgcaacattg	tcgtcacacc	gtagatatgg	cccaagcttt	gagatctgct	3480
cgagaaaagg	gactcatcga	cgacaaaacc	ctcgtcattg	agctgggacc	gaagccatta	3540
atctcgggca	tggtgaaaat	gacactggga	gacaaaatta	gcaccttacc	cactctagca	3600
cctaacaagg	ccattttggc	cagcctgcag	aagattctca	cctcggctca	cacgggtggg	3660
tggtgatatta	attggaagaa	atatcacgcc	cctttcgcc	cctcccagaa	ggtgggtggat	3720
ctgcgagcgt	acggctggga	tttgaaggac	tactacatcc	cgtatcaggg	tgactgggtgt	3780
ctgcacgcgc	accagcagga	ttgtaagtgc	gccgctcctg	gccacgaaat	caaaacggcc	3840
gactaccaag	tgccctcctga	gtcgacgcct	caccgtccat	ccaagctgga	ccctagcaag	3900
gaggcccttcc	ccgaaatcaa	gaccaccacg	acactccatc	gagtgggtgga	agagacgacc	3960
aaaccttttg	gcgccaccct	agttgtggag	acagacatat	ctcggaagga	tgtcaacggc	4020
ctcgctcgag	ggcaccttgt	cgtatgggatc	cctttgtgta	ccccctcctt	ttatgctgac	4080
atcgccatgc	aagtgggcca	atacagtatg	caacggctcc	gtgcgggaca	tcggggggcc	4140
ggtgccatag	atggccttgt	ggacgtgtcc	gacatgggtg	tggaacaaag	ggttggttccc	4200
catgggaagg	gacctcaatt	gctgcgcacg	acgcttacca	tgagtggtgc	gcccagggt	4260
gctgctacta	cggaagcgc	caaagtcaaa	ttcgccacct	atthttgccga	tggaaggtcc	4320
gatacggagc	atgccagctg	tactgtcaga	ttcacaagcg	atgcacagtt	gaaatctcta	4380
cgccggctctg	tgctcgagta	caagaccac	attcgtcagt	tacatgatgg	ccatgctaag	4440
ggacagttca	tgcgatacaa	taggaagacc	gggtacaagc	tcatagagcag	catggctcgg	4500
tttaatcccg	actacatgct	cctagattat	ctgggtgctga	acgaagcaga	gaacgaggca	4560
gcaagtgggtg	tagacttctc	gttgggatcg	tcggaaggca	ccttcgcagc	tcaccacagct	4620

cacgtcgatg ccatactca ggtggccggc tttgctatga atgccaatga caatgtcgac 4680  
 attgagaaac aggtctacgt taatcacggt tgggactcgt tccagatcta ccaaccgctg 4740  
 gataatagca agtcttacca ggtctacacc aagatgggtc aagcgaagga gaatgatttg 4800  
 gtgcatggcg atgtggtagt tctggacgga gaacaaatcg ttgctttctt ccgcgccctt 4860  
 acgctgcgat cagttcctcg tgggtgactg cgtgtcgtcc tgcagactac agtgaaaaag 4920  
 gccgatcgcc aactaggatt taagacaatg ccgtcgccgc cggcccccgc aacgacaatg 4980  
 ccaatatcgc cttataaacc agctaatact caggtttcca gccaaagctat tccagcagag 5040  
 gccactcatt ctcacacccc gccacagcca aagcattccc cgggtaccgga aactgccgga 5100  
 agcgtccag cggcaaaaagg agtaggtgtc agtaacgaaa agttagatgc tgtaatgcga 5160  
 gtcgtttcgg aggagagtgg aattgccctc gaggagctca ccgatgacag caactttgct 5220  
 gacatgggca tgcactctct gagttcaatg gtcattggga gccgcttcag agaggacctg 5280  
 gggctggacc tggggcctga gttttctctt ttcattgact gcactaccgt gcgtgccttg 5340  
 aaagacttca tgttgggaag cggggatgct ggcagtggct ccaatgtaga agatcctccc 5400  
 ccatacagcta ctcccggcat caacccega accgattggt ctagcagtgc ctctgatagt 5460  
 attttcgcca gcgaagacca cggtcattcg agtgagtccg gcgccgacac cggaagtccg 5520  
 cctgcacttg atctgaagcc ctactgccgc cctcaactt ctgtcgtcct acaaggtcta 5580  
 cctatggtgg cgcggaaaac tctgtttatg ctccctgatg gcgggggggc tgcgttctcc 5640  
 taagcctccc tgccgcgcct caaatcagat actgccgttg tgggcctgaa ttgcccttat 5700  
 getcgggac ccgagaacat gaactgcaca catggagcta tgattgagag cttttgcaat 5760  
 gagatccggc ggcgacagcc acggggcccc tatcacctgg gcggctgggc gtccggtggt 5820  
 gcattcgctt acgtcgtggc cgaggcactt gttaaccaag gcgaggaggt gcattcgta 5880  
 atcatcattg atgcgcctat tccccaaagc atggaacaac tccccgagc attttacgag 5940  
 cactgcaata gcattggatt gttcgctacc cagccggggg ctagtccgga cggctcgact 6000  
 gagcctccat cctacttaat cccacacttt accgctgtgg tggatgtgat gctggattac 6060  
 aagctggccc cgttgcatgc gcgccggatg cccaaggtcg gcatcgtctg ggccggcagat 6120  
 acagtcatgg acgagcggga cgctcccaag atgaaaggaa tgcattttat gattcagaag 6180  
 cggacggaat ttggtcccga tgggtgggat acgatcatgc ccggggcctc gtttgacatt 6240  
 gtccgagcag acggtgctaa tcattttacg ttgatgcaaa aggaacatgt ctctataatt 6300  
 agcgatctga tcgaccgggt catggcttag 6330

&lt;210&gt; 12

&lt;211&gt; 2109

&lt;212&gt; PRT

&lt;213&gt; Aspergillus parasiticus

&lt;400&gt; 12

Met Ala Gln Ser Arg Gln Leu Phe Leu Phe Gly Asp Gln Thr Ala Asp  
 1 5 10 15  
 Phe Val Pro Lys Leu Arg Ser Leu Leu Ser Val Gln Asp Ser Pro Ile  
 20 25 30  
 Leu Ala Ala Phe Leu Asp Gln Ser His Tyr Val Val Arg Ala Gln Met  
 35 40 45  
 Leu Gln Ser Met Asn Thr Val Asp His Lys Leu Ala Arg Thr Ala Asp  
 50 55 60  
 Leu Arg Gln Met Val Gln Lys Tyr Val Asp Gly Lys Leu Thr Pro Ala  
 65 70 75 80  
 Phe Arg Thr Ala Leu Val Cys Leu Cys Gln Leu Gly Cys Phe Ile Arg  
 85 90 95

40

Glu Tyr Glu Glu Ser Gly Asn Met Tyr Pro Gln Pro Ser Asp Ser Tyr	100	105	110
Val Leu Gly Phe Cys Met Gly Ser Leu Ala Ala Val Ala Val Ser Cys	115	120	125
Ser Arg Ser Leu Ser Glu Leu Leu Pro Ile Ala Val Gln Thr Val Leu	130	135	140
Ile Ala Phe Arg Leu Gly Leu Cys Ala Leu Glu Met Arg Asp Arg Val	145	150	155
Asp Gly Cys Ser Asp Asp Arg Gly Asp Pro Trp Ser Thr Ile Val Trp	165	170	175
Gly Leu Asp Pro Gln Gln Ala Arg Asp Gln Ile Glu Val Phe Cys Arg	180	185	190
Thr Thr Asn Val Pro Gln Thr Arg Arg Pro Trp Ile Ser Cys Ile Ser	195	200	205
Lys Asn Ala Ile Thr Leu Ser Gly Ser Pro Ser Thr Leu Arg Ala Phe	210	215	220
Cys Ala Met Pro Gln Met Ala Gln His Arg Thr Ala Pro Ile Pro Ile	225	230	235
Cys Leu Pro Ala His Asn Gly Ala Leu Phe Thr Gln Ala Asp Ile Thr	245	250	255
Thr Ile Leu Asp Thr Thr Pro Thr Thr Pro Trp Glu Gln Leu Pro Gly	260	265	270
Gln Ile Pro Tyr Ile Ser His Val Thr Gly Asn Val Val Gln Thr Ser	275	280	285
Asn Tyr Arg Asp Leu Ile Glu Val Ala Leu Ser Glu Thr Leu Leu Glu	290	295	300
Gln Val Arg Leu Asp Leu Val Glu Thr Gly Leu Pro Arg Leu Leu Gln	305	310	315
Ser Arg Gln Val Lys Ser Val Thr Ile Val Pro Phe Leu Thr Arg Met	325	330	335
Asn Glu Thr Met Ser Asn Ile Leu Pro Asp Ser Phe Ile Ser Thr Glu	340	345	350
Thr Arg Thr Asp Thr Gly Arg Ala Ile Pro Ala Ser Gly Arg Pro Gly	355	360	365
Ala Gly Lys Cys Lys Leu Ala Ile Val Ser Met Ser Gly Arg Phe Pro	370	375	380
Glu Ser Pro Thr Thr Glu Ser Phe Trp Asp Leu Leu Tyr Lys Gly Leu	385	390	395
			400

Asp	Val	Cys	Lys	Glu	Val	Pro	Arg	Arg	Arg	Trp	Asp	Ile	Asn	Thr	His	
				405						410						415
Val	Asp	Pro	Ser	Gly	Lys	Ala	Arg	Asn	Lys	Gly	Ala	Thr	Lys	Trp	Gly	
			420					425					430			
Cys	Trp	Leu	Asp	Phe	Ser	Gly	Asp	Phe	Asp	Pro	Arg	Phe	Phe	Gly	Ile	
		435					440					445				
Ser	Pro	Lys	Glu	Ala	Pro	Gln	Met	Asp	Pro	Ala	Gln	Arg	Met	Ala	Leu	
		450				455					460					
Met	Ser	Thr	Tyr	Glu	Ala	Met	Glu	Arg	Ala	Gly	Leu	Val	Pro	Asp	Thr	
465					470					475					480	
Thr	Pro	Ser	Thr	Gln	Arg	Asp	Arg	Ile	Gly	Val	Phe	His	Gly	Val	Thr	
				485					490					495		
Ser	Asn	Asp	Trp	Met	Glu	Thr	Asn	Thr	Ala	Gln	Asn	Ile	Asp	Thr	Tyr	
			500					505					510			
Phe	Ile	Thr	Gly	Gly	Asn	Arg	Gly	Phe	Ile	Pro	Gly	Arg	Ile	Asn	Phe	
		515					520					525				
Cys	Phe	Glu	Phe	Ala	Gly	Pro	Ser	Tyr	Thr	Asn	Asp	Thr	Ala	Cys	Ser	
		530				535						540				
Ser	Ser	Leu	Ala	Ala	Ile	His	Leu	Ala	Cys	Asn	Ser	Leu	Trp	Arg	Gly	
545					550					555					560	
Asp	Cys	Asp	Thr	Ala	Val	Ala	Gly	Gly	Thr	Asn	Met	Ile	Tyr	Thr	Pro	
				565					570					575		
Asp	Gly	His	Thr	Gly	Leu	Asp	Lys	Gly	Phe	Phe	Leu	Ser	Arg	Thr	Gly	
			580					585					590			
Asn	Cys	Lys	Pro	Tyr	Asp	Asp	Lys	Ala	Asp	Gly	Tyr	Cys	Arg	Ala	Glu	
		595					600					605				
Gly	Val	Gly	Thr	Val	Phe	Ile	Lys	Arg	Leu	Glu	Asp	Ala	Leu	Ala	Asp	
		610				615					620					
Asn	Asp	Pro	Ile	Leu	Gly	Val	Ile	Leu	Asp	Ala	Lys	Thr	Asn	His	Ser	
625				630						635					640	
Ala	Met	Ser	Glu	Ser	Met	Thr	Arg	Pro	His	Val	Gly	Ala	Gln	Ile	Asp	
				645					650					655		
Asn	Met	Thr	Ala	Ala	Leu	Asn	Thr	Thr	Gly	Leu	His	Pro	Asn	Asp	Phe	
			660					665					670			
Ser	Tyr	Ile	Glu	Met	His	Gly	Thr	Gly	Thr	Gln	Val	Gly	Asp	Ala	Val	
		675					680					685				
Glu	Met	Glu	Ser	Val	Leu	Ser	Val	Phe	Ala	Pro	Ser	Glu	Thr	Ala	Arg	
		690				695						700				

42

Lys Ala Asp Gln Pro Leu Phe Val Gly Ser Ala Lys Ala Asn Val Gly  
705 710 715 720

His Gly Glu Gly Val Ser Gly Val Thr Ser Leu Ile Lys Val Leu Met  
725 730 735

Met Met Gln His Asp Thr Ile Pro Pro His Cys Gly Ile Lys Pro Gly  
740 745 750

Ser Lys Ile Asn Arg Asn Phe Pro Asp Leu Gly Ala Arg Asn Val His  
755 760 765

Ile Ala Phe Glu Pro Lys Pro Trp Pro Arg Thr His Thr Pro Arg Arg  
770 775 780

Val Leu Ile Asn Asn Phe Ser Ala Ala Gly Gly Asn Thr Ala Leu Ile  
785 790 795 800

Val Glu Asp Ala Pro Glu Arg His Trp Pro Thr Glu Lys Asp Pro Arg  
805 810 815

Ser Ser His Ile Val Ala Leu Ser Ala His Val Gly Ala Ser Met Lys  
820 825 830

Thr Asn Leu Glu Arg Leu His Gln Tyr Leu Leu Lys Asn Pro His Thr  
835 840 845

Asp Leu Ala Gln Leu Ser Tyr Thr Thr Thr Ala Arg Arg Trp His Tyr  
850 855 860

Leu His Arg Val Ser Val Thr Gly Ala Ser Val Glu Glu Val Thr Arg  
865 870 875 880

Lys Leu Glu Met Ala Ile Gln Asn Gly Asp Gly Val Ser Arg Pro Lys  
885 890 895

Ser Lys Pro Lys Ile Leu Phe Ala Phe Thr Gly Gln Gly Ser Gln Tyr  
900 905 910

Ala Thr Met Gly Lys Gln Val Tyr Asp Ala Tyr Pro Ser Phe Arg Glu  
915 920 925

Asp Leu Glu Lys Phe Asp Arg Leu Ala Gln Ser His Gly Phe Pro Ser  
930 935 940

Phe Leu His Val Cys Thr Ser Pro Lys Gly Asp Val Glu Glu Met Ala  
945 950 955 960

Pro Val Val Val Gln Leu Ala Ile Thr Cys Leu Gln Met Ala Leu Thr  
965 970 975

Asn Leu Met Thr Ser Phe Gly Ile Arg Pro Asp Val Thr Val Gly His  
980 985 990

Ser Leu Gly Glu Phe Ala Ala Leu Tyr Ala Ala Gly Val Leu Ser Ala  
995 1000 1005

43

Ser: Asp Val Val Tyr Leu Val Gly Gln Arg Ala Glu Leu Leu Gln Glu  
 1010 1015 1020  
 Arg Cys Gln Arg Gly Thr His Ala Met Leu Ala Val Lys Ala Thr Pro  
 1025 1030 1035 1040  
 Glu Ala Leu Ser Gln Trp Ile Gln Asp His Asp Cys Glu Val Ala Cys  
 1045 1050 1055  
 Ile Asn Gly Pro Glu Asp Thr Val Leu Ser Gly Thr Thr Lys Asn Val  
 1060 1065 1070  
 Ala Glu Val Gln Arg Ala Met Thr Asp Asn Gly Ile Lys Cys Thr Leu  
 1075 1080 1085  
 Leu Lys Leu Pro Phe Ala Phe His Ser Ala Gln Val Gln Pro Ile Leu  
 1090 1095 1100  
 Asp Asp Phe Glu Ala Leu Ala Gln Gly Ala Thr Phe Ala Lys Pro Gln  
 1105 1110 1115 1120  
 Leu Leu Ile Leu Ser Pro Leu Leu Arg Thr Glu Ile His Glu Gln Gly  
 1125 1130 1135  
 Val Val Thr Pro Ser Tyr Val Ala Gln His Cys Arg His Thr Val Asp  
 1140 1145 1150  
 Met Ala Gln Ala Leu Arg Ser Ala Arg Glu Lys Gly Leu Ile Asp Asp  
 1155 1160 1165  
 Lys Thr Leu Val Ile Glu Leu Gly Pro Lys Pro Leu Ile Ser Gly Met  
 1170 1175 1180  
 Val Lys Met Thr Leu Gly Asp Lys Ile Ser Thr Leu Pro Thr Leu Ala  
 1185 1190 1195 1200  
 Pro Asn Lys Ala Ile Trp Pro Ser Leu Gln Lys Ile Leu Thr Ser Val  
 1205 1210 1215  
 Tyr Thr Gly Gly Trp Asp Ile Asn Trp Lys Lys Tyr His Ala Pro Phe  
 1220 1225 1230  
 Ala Ser Ser Gln Lys Val Val Asp Leu Pro Ser Tyr Gly Trp Asp Leu  
 1235 1240 1245  
 Lys Asp Tyr Tyr Ile Pro Tyr Gln Gly Asp Trp Cys Leu His Arg His  
 1250 1255 1260  
 Gln Gln Asp Cys Lys Cys Ala Ala Pro Gly His Glu Ile Lys Thr Ala  
 1265 1270 1275 1280  
 Asp Tyr Gln Val Pro Pro Glu Ser Thr Pro His Arg Pro Ser Lys Leu  
 1285 1290 1295  
 Asp Pro Ser Lys Glu Ala Phe Pro Glu Ile Lys Thr Thr Thr Thr Leu  
 1300 1305 1310

## 44

His Arg Val Val Glu Glu Thr Thr Lys Pro Leu Gly Ala Thr Leu Val  
 1315 1320 1325

Val Glu Thr Asp Ile Ser Arg Lys Asp Val Asn Gly Leu Ala Arg Gly  
 1330 1335 1340

His Leu Val Asp Gly Ile Pro Leu Cys Thr Pro Ser Phe Tyr Ala Asp  
 1345 1350 1355 1360

Ile Ala Met Gln Val Gly Gln Tyr Ser Met Gln Arg Leu Arg Ala Gly  
 1365 1370 1375

His Pro Gly Ala Gly Ala Ile Asp Gly Leu Val Asp Val Ser Asp Met  
 1380 1385 1390

Val Val Asp Lys Ala Leu Val Pro His Gly Lys Gly Pro Gln Leu Leu  
 1395 1400 1405

Arg Thr Thr Leu Thr Met Glu Trp Pro Pro Lys Ala Ala Ala Thr Thr  
 1410 1415 1420

Arg Ser Ala Lys Val Lys Phe Ala Thr Tyr Phe Ala Asp Gly Lys Leu  
 1425 1430 1435 1440

Asp Thr Glu His Ala Ser Cys Thr Val Arg Phe Thr Ser Asp Ala Gln  
 1445 1450 1455

Leu Lys Ser Leu Arg Arg Ser Val Ser Glu Tyr Lys Thr His Ile Arg  
 1460 1465 1470

Gln Leu His Asp Gly His Ala Lys Gly Gln Phe Met Arg Tyr Asn Arg  
 1475 1480 1485

Lys Thr Gly Tyr Lys Leu Met Ser Ser Met Ala Arg Phe Asn Pro Asp  
 1490 1495 1500

Tyr Met Leu Leu Asp Tyr Leu Val Leu Asn Glu Ala Glu Asn Glu Ala  
 1505 1510 1515 1520

Ala Ser Gly Val Asp Phe Ser Leu Gly Ser Ser Glu Gly Thr Phe Ala  
 1525 1530 1535

Ala His Pro Ala His Val Asp Ala Ile Thr Gln Val Ala Gly Phe Ala  
 1540 1545 1550

Met Asn Ala Asn Asp Asn Val Asp Ile Glu Lys Gln Val Tyr Val Asn  
 1555 1560 1565

His Gly Trp Asp Ser Phe Gln Ile Tyr Gln Pro Leu Asp Asn Ser Lys  
 1570 1575 1580

Ser Tyr Gln Val Tyr Thr Lys Met Gly Gln Ala Lys Glu Asn Asp Leu  
 1585 1590 1595 1600

Val His Gly Asp Val Val Val Leu Asp Gly Glu Gln Ile Val Ala Phe  
 1605 1610 1615

45

Phe Arg Gly Leu Thr Leu Arg Ser Val Pro Arg Gly Ala Leu Arg Val		
1620	1625	1630
Val Leu Gln Thr Thr Val Lys Lys Ala Asp Arg Gln Leu Gly Phe Lys		
1635	1640	1645
Thr Met Pro Ser Pro Pro Pro Pro Thr Thr Thr Met Pro Ile Ser Pro		
1650	1655	1660
Tyr Lys Pro Ala Asn Thr Gln Val Ser Ser Gln Ala Ile Pro Ala Glu		
1665	1670	1675 1680
Ala Thr His Ser His Thr Pro Pro Gln Pro Lys His Ser Pro Val Pro		
1685	1690	1695
Glu Thr Ala Gly Ser Ala Pro Ala Ala Lys Gly Val Gly Val Ser Asn		
1700	1705	1710
Glu Lys Leu Asp Ala Val Met Arg Val Val Ser Glu Glu Ser Gly Ile		
1715	1720	1725
Ala Leu Glu Glu Leu Thr Asp Asp Ser Asn Phe Ala Asp Met Gly Ile		
1730	1735	1740
Asp Ser Leu Ser Ser Met Val Ile Gly Ser Arg Phe Arg Glu Asp Leu		
1745	1750	1755 1760
Gly Leu Asp Leu Gly Pro Glu Phe Ser Leu Phe Ile Asp Cys Thr Thr		
1765	1770	1775
Val Arg Ala Leu Lys Asp Phe Met Leu Gly Ser Gly Asp Ala Gly Ser		
1780	1785	1790
Gly Ser Asn Val Glu Asp Pro Pro Pro Ser Ala Thr Pro Gly Ile Asn		
1795	1800	1805
Pro Glu Thr Asp Trp Ser Ser Ser Ala Ser Asp Ser Ile Phe Ala Ser		
1810	1815	1820
Glu Asp His Gly His Ser Ser Glu Ser Gly Ala Asp Thr Gly Ser Pro		
1825	1830	1835 1840
Pro Ala Leu Asp Leu Lys Pro Tyr Cys Arg Pro Ser Thr Ser Val Val		
1845	1850	1855
Leu Gln Gly Leu Pro Met Val Ala Arg Lys Thr Leu Phe Met Leu Pro		
1860	1865	1870
Asp Gly Gly Gly Ser Ala Phe Ser Tyr Ala Ser Leu Pro Arg Leu Lys		
1875	1880	1885
Ser Asp Thr Ala Val Val Gly Leu Asn Cys Pro Tyr Ala Arg Asp Pro		
1890	1895	1900
Glu Asn Met Asn Cys Thr His Gly Ala Met Ile Glu Ser Phe Cys Asn		
1905	1910	1915 1920

46

Glu Ile Arg Arg Arg Gln Pro Arg Gly Pro Tyr His Leu Gly Gly Trp  
 1925 1930 1935

Ser Ser Gly Gly Ala Phe Ala Tyr Val Val Ala Glu Ala Leu Val Asn  
 1940 1945 1950

Gln Gly Glu Glu Val His Ser Leu Ile Ile Ile Asp Ala Pro Ile Pro  
 1955 1960 1965

Gln Ala Met Glu Gln Leu Pro Arg Ala Phe Tyr Glu His Cys Asn Ser  
 1970 1975 1980

Ile Gly Leu Phe Ala Thr Gln Pro Gly Ala Ser Pro Asp Gly Ser Thr  
 1985 1990 1995 2000

Glu Pro Pro Ser Tyr Leu Ile Pro His Phe Thr Ala Val Val Asp Val  
 2005 2010 2015

Met Leu Asp Tyr Lys Leu Ala Pro Leu His Ala Arg Arg Met Pro Lys  
 2020 2025 2030

Val Gly Ile Val Trp Ala Ala Asp Thr Val Met Asp Glu Arg Asp Ala  
 2035 2040 2045

Pro Lys Met Lys Gly Met His Phe Met Ile Gln Lys Arg Thr Glu Phe  
 2050 2055 2060

Gly Pro Asp Gly Trp Asp Thr Ile Met Pro Gly Ala Ser Phe Asp Ile  
 2065 2070 2075 2080

Val Arg Ala Asp Gly Ala Asn His Phe Thr Leu Met Gln Lys Glu His  
 2085 2090 2095

Val Ser Ile Ile Ser Asp Leu Ile Asp Arg Val Met Ala  
 2100 2105

&lt;210&gt; 13

&lt;211&gt; 1986

&lt;212&gt; PRT

&lt;213&gt; Aspergillus nidulans

&lt;400&gt; 13

Met Glu Asp Pro Tyr Arg Val Tyr Leu Phe Gly Asp Gln Thr Gly Asp  
 1 5 10 15

Phe Glu Val Gly Leu Arg Arg Leu Leu Gln Ala Lys Asn His Ser Leu  
 20 25 30

Leu Ser Ser Phe Leu Gln Arg Ser Tyr His Ala Val Arg Gln Glu Ile  
 35 40 45

Ser His Leu Pro Pro Ser Glu Arg Ser Thr Phe Pro Arg Phe Thr Ser  
 50 55 60

Ile Gly Asp Leu Leu Ala Arg His Cys Glu Ser Pro Gly Asn Pro Ala  
 65 70 75 80

47

Ile Glu Ser Val Leu Thr Cys Ile Tyr Gln Leu Gly Cys Phe Ile Asn  
85 90 95  
Tyr Tyr Gly Asp Leu Gly His Thr Phe Pro Ser His Ser Gln Ser Gln  
100 105 110  
Leu Val Gly Leu Cys Thr Gly Leu Leu Ser Cys Ala Ala Val Ser Cys  
115 120 125  
Ala Ser Asn Ile Gly Glu Leu Leu Lys Pro Ala Val Glu Val Val Val  
130 135 140  
Val Ala Leu Arg Leu Gly Leu Cys Val Tyr Arg Val Arg Lys Leu Phe  
145 150 155 160  
Gly Gln Asp Gln Ala Ala Pro Leu Ser Trp Ser Ala Leu Val Ser Gly  
165 170 175  
Leu Ser Glu Ser Glu Gly Thr Ser Leu Ile Asp Lys Phe Thr Arg Arg  
180 185 190  
Asn Val Ile Pro Pro Ser Ser Arg Pro Tyr Ile Ser Ala Val Cys Ala  
195 200 205  
Asn Thr Leu Thr Ile Ser Gly Pro Pro Val Val Leu Asn Gln Phe Leu  
210 215 220  
Asp Thr Phe Ile Ser Gly Lys Asn Lys Ala Val Met Val Pro Ile His  
225 230 235 240  
Gly Pro Phe His Ala Ser His Leu Tyr Glu Lys Arg Asp Val Glu Trp  
245 250 255  
Ile Leu Lys Ser Cys Asn Val Glu Thr Ile Arg Asn His Lys Pro Arg  
260 265 270  
Ile Pro Val Leu Ser Ser Asn Thr Gly Glu Leu Ile Val Val Glu Asn  
275 280 285  
Met Glu Gly Phe Leu Lys Ile Ala Leu Glu Glu Ile Leu Leu Arg Gln  
290 295 300  
Met Ser Trp Asp Lys Val Thr Asp Ser Cys Ile Ser Ile Leu Lys Ser  
305 310 315 320  
Val Gly Asp Asn Lys Pro Lys Lys Leu Leu Pro Ile Ser Ser Thr Ala  
325 330 335  
Thr Gln Ser Leu Phe Asn Ser Leu Lys Lys Ser Asn Leu Val Asn Ile  
340 345 350  
Glu Val Asp Gly Gly Ile Ser Asp Phe Ala Ala Glu Thr Gln Leu Val  
355 360 365  
Asn Gln Thr Gly Arg Ala Glu Leu Ser Lys Ile Ala Ile Ile Gly Met  
370 375 380

48

Ser Gly Arg Phe Pro Glu Ala Asp Ser Pro Gln Asp Phe Trp Asn Leu  
385 390 395 400  
Leu Tyr Lys Gly Leu Asp Val His Arg Lys Val Pro Glu Asp Arg Trp  
405 410 415  
Asp Ala Asp Ala His Val Asp Leu Thr Gly Thr Ala Thr Asn Thr Ser  
420 425 430  
Lys Val Pro Tyr Gly Cys Trp Ile Arg Glu Pro Gly Leu Phe Asp Pro  
435 440 445  
Arg Phe Phe Asn Met Ser Pro Arg Glu Ala Leu Gln Ala Asp Pro Ala  
450 455 460  
Gln Arg Leu Ala Leu Leu Thr Ala Tyr Glu Ala Leu Glu Gly Ala Gly  
465 470 475 480  
Phe Val Pro Asp Ser Thr Pro Ser Thr Gln Arg Asp Arg Val Gly Ile  
485 490 495  
Phe Tyr Gly Met Thr Ser Asp Asp Tyr Arg Glu Val Asn Ser Gly Gln  
500 505 510  
Asp Ile Asp Thr Tyr Phe Ile Pro Gly Gly Asn Arg Ala Phe Thr Pro  
515 520 525  
Gly Arg Ile Asn Tyr Tyr Phe Lys Phe Ser Gly Pro Ser Val Ser Val  
530 535 540  
Asp Thr Ala Cys Ser Ser Ser Leu Ala Ala Ile His Leu Ala Cys Asn  
545 550 555 560  
Ser Ile Trp Arg Asn Asp Cys Asp Thr Ala Ile Thr Gly Gly Val Asn  
565 570 575  
Ile Leu Thr Asn Pro Asp Asn His Ala Gly Leu Asp Arg Gly His Phe  
580 585 590  
Leu Ser Arg Thr Gly Asn Cys Asn Thr Phe Asp Asp Gly Ala Asp Gly  
595 600 605  
Tyr Cys Arg Ala Asp Gly Val Gly Thr Val Val Leu Lys Arg Leu Glu  
610 615 620  
Asp Ala Leu Ala Asp Asn Asp Pro Ile Leu Gly Val Ile Asn Gly Ala  
625 630 635 640  
Tyr Thr Asn His Ser Ala Glu Ala Val Ser Ile Thr Arg Pro His Val  
645 650 655  
Gly Ala Gln Ala Phe Ile Phe Lys Lys Leu Leu Asn Glu Ala Asn Val  
660 665 670  
Asp Pro Lys Asn Ile Ser Tyr Ile Glu Met His Gly Thr Gly Thr Gln  
675 680 685

## 49

Ala Gly Asp Ala Val Glu Met Gln Ser Val Leu Asp Val Phe Ala Pro	690	695	700
Asp His Arg Arg Gly Pro Gly Gln Ser Leu His Leu Gly Ser Ala Lys	705	710	715 720
Ser Asn Ile Gly His Gly Glu Ser Ala Ser Gly Val Thr Ser Leu Val	725	730	735
Lys Val Leu Leu Met Met Lys Glu Asn Met Ile Pro Pro His Cys Gly	740	745	750
Ile Lys Thr Lys Ile Asn His Asn Phe Pro Thr Asp Leu Ala Gln Arg	755	760	765
Asn Val His Ile Ala Leu Gln Pro Thr Ala Trp Asn Arg Pro Ser Phe	770	775	780
Gly Lys Arg Gln Ile Phe Leu Asn Asn Phe Ser Ala Ala Gly Gly Asn	785	790	795 800
Thr Ala Leu Leu Leu Glu Asp Gly Pro Val Ser Asp Pro Glu Gly Glu	805	810	815
Asp Lys Arg Arg Thr His Val Ile Thr Leu Ser Ala Arg Ser Gln Thr	820	825	830
Ala Leu Gln Asn Asn Ile Asp Ala Leu Cys Gln Tyr Ile Ser Glu Gln	835	840	845
Glu Lys Thr Phe Gly Val Lys Asp Ser Asn Ala Leu Pro Ser Leu Ala	850	855	860
Tyr Thr Thr Thr Ala Arg Arg Ile His His Pro Phe Arg Val Thr Ala	865	870	875 880
Ile Gly Ser Ser Phe Gln Glu Met Arg Asp Ser Leu Ile Ala Ser Ser	885	890	895
Arg Lys Glu Phe Val Ala Val Pro Ala Lys Thr Pro Gly Ile Gly Phe	900	905	910
Leu Phe Thr Gly Gln Gly Ala Gln Tyr Ala Ala Met Gly Lys Gln Leu	915	920	925
Tyr Glu Asp Cys Ser His Phe Arg Ser Ala Ile Glu His Leu Asp Cys	930	935	940
Ile Ser Gln Gly Gln Asp Leu Pro Ser Ile Leu Pro Leu Val Asp Gly	945	950	955 960
Ser Leu Pro Leu Ser Glu Leu Ser Pro Val Val Val Gln Leu Gly Thr	965	970	975
Thr Cys Val Gln Met Ala Leu Ser Ser Phe Trp Ala Ser Leu Gly Ile	980	985	990

## 50

Thr Pro Ser Phe Val Leu Gly His Ser Leu Gly Asp Phe Ala Ala Met		
995	1000	1005
Asn Ala Ala Gly Val Leu Ser Thr Ser Asp Thr Ile Tyr Ala Cys Gly		
1010	1015	1020
Arg Arg Ala Gln Leu Leu Thr Glu Arg Cys Gln Pro Gly Thr His Ala		
1025	1030	1035 1040
Met Leu Ala Ile Lys Ala Pro Leu Val Glu Val Lys Gln Leu Leu Asn		
1045	1050	1055
Glu Lys Val His Asp Met Ala Cys Ile Asn Ser Pro Ser Glu Thr Val		
1060	1065	1070
Ile Ser Gly Pro Lys Ser Ser Ile Asp Glu Leu Ser Arg Ala Cys Ser		
1075	1080	1085
Glu Lys Gly Leu Lys Ser Thr Ile Leu Thr Val Pro Tyr Ala Phe His		
1090	1095	1100
Ser Ala Gln Val Glu Pro Ile Leu Glu Asp Leu Glu Lys Ala Leu Gln		
1105	1110	1115 1120
Gly Ile Thr Phe Asn Lys Pro Ser Val Pro Phe Val Ser Ala Leu Leu		
1125	1130	1135
Gly Glu Val Ile Thr Glu Ala Gly Ser Asn Ile Leu Asn Ala Glu Tyr		
1140	1145	1150
Leu Val Arg His Cys Arg Glu Thr Val Asn Phe Leu Ser Ala Phe Glu		
1155	1160	1165
Ala Val Arg Asn Ala Lys Leu Gly Gly Asp Gln Thr Leu Trp Leu Glu		
1170	1175	1180
Val Gly Pro His Thr Val Cys Ser Gly Met Val Lys Ala Thr Leu Gly		
1185	1190	1195 1200
Pro Gln Thr Thr Thr Met Ala Ser Leu Arg Arg Asp Glu Asp Thr Trp		
1205	1210	1215
Lys Val Leu Ser Asn Ser Leu Ser Ser Leu Tyr Leu Ala Gly Val Asp		
1220	1225	1230
Ile Asn Trp Lys Gln Tyr His Gln Asp Phe Ser Ser Ser His Arg Val		
1235	1240	1245
Leu Pro Leu Pro Thr Tyr Lys Trp Asp Leu Lys Asn Tyr Trp Ile Pro		
1250	1255	1260
Tyr Arg Asn Asn Phe Cys Leu Thr Lys Gly Ser Ser Met Ser Ala Ala		
1265	1270	1275 1280
Ser Ala Ser Leu Gln Pro Thr Phe Leu Thr Thr Ser Ala Gln Arg Val		
1285	1290	1295

## 51

Val Glu Ser Arg Asp Asp Gly Leu Thr Ala Thr Val Val Val His Asn  
 1300 1305 1310  
 Asp Ile Ala Asp Pro Asp Leu Asn Arg Val Ile Gln Gly His Lys Val  
 1315 1320 1325  
 Asn Gly Ala Ala Leu Cys Pro Ser Ser Leu Tyr Ala Asp Ser Ala Gln  
 1330 1335 1340  
 Thr Leu Ala Glu Tyr Leu Ile Glu Lys Tyr Lys Pro Glu Leu Lys Gly  
 1345 1350 1355 1360  
 Ser Gly Leu Asp Val Cys Asn Val Thr Val Pro Lys Pro Leu Ile Ala  
 1365 1370 1375  
 Lys Thr Gly Lys Glu Gln Phe Arg Ile Ser Ala Thr Ala Asn Trp Val  
 1380 1385 1390  
 Asp Lys His Val Ser Val Gln Val Phe Ser Val Thr Ala Glu Gly Lys  
 1395 1400 1405  
 Lys Leu Ile Asp His Ala His Cys Glu Val Lys Leu Phe Asp Cys Met  
 1410 1415 1420  
 Ala Ala Asp Leu Glu Trp Lys Arg Gly Ser Tyr Leu Val Lys Arg Ser  
 1425 1430 1435 1440  
 Ile Glu Leu Leu Glu Asn Ser Ala Val Lys Gly Asp Ala His Arg Leu  
 1445 1450 1455  
 Arg Arg Gly Met Val Tyr Lys Leu Phe Ser Ala Leu Val Asp Tyr Asp  
 1460 1465 1470  
 Glu Asn Tyr Gln Ser Ile Arg Glu Val Ile Leu Asp Ser Glu His His  
 1475 1480 1485  
 Glu Ala Thr Ala Leu Val Lys Phe Gln Ala Pro Gln Ala Asn Phe His  
 1490 1495 1500  
 Arg Asn Pro Tyr Trp Ile Asp Ser Phe Gly His Leu Ser Gly Phe Ile  
 1505 1510 1515 1520  
 Met Asn Ala Ser Asp Gly Thr Asp Ser Lys Ser Gln Val Phe Val Asn  
 1525 1530 1535  
 His Gly Trp Asp Ser Met Arg Cys Leu Lys Lys Phe Ser Ala Asp Val  
 1540 1545 1550  
 Thr Tyr Arg Thr Tyr Val Arg Met Gln Pro Trp Arg Asp Ser Ile Trp  
 1555 1560 1565  
 Ala Gly Asn Val Tyr Ile Phe Glu Gly Asp Asp Ile Ile Ala Val Phe  
 1570 1575 1580  
 Gly Gly Val Lys Phe Gln Ala Leu Ser Arg Lys Ile Leu Asp Ile Ala  
 1585 1590 1595 1600

## 52

Leu Pro Pro Ala Gly Leu Ser Lys Ala Gln Thr Ser Pro Ile Gln Ser  
 1605 1610 1615  
 Ser Ala Pro Gln Lys Pro Ile Glu Thr Ala Lys Pro Thr Ser Arg Pro  
 1620 1625 1630  
 Ala Pro Pro Val Thr Met Lys Ser Phe Val Lys Lys Ser Ala Gly Pro  
 1635 1640 1645  
 Ser Val Val Val Arg Ala Leu Asn Ile Leu Ala Ser Glu Val Gly Leu  
 1650 1655 1660  
 Ser Glu Ser Asp Met Ser Asp Asp Leu Val Phe Ala Asp Tyr Gly Val  
 1665 1670 1675 1680  
 Asp Ser Leu Leu Ser Leu Thr Val Thr Gly Lys Tyr Arg Glu Glu Leu  
 1685 1690 1695  
 Asn Leu Asp Met Asp Ser Ser Val Phe Ile Glu His Pro Thr Val Gly  
 1700 1705 1710  
 Asp Phe Lys Arg Phe Val Thr Gln Leu Ser Pro Ser Val Ala Ser Asp  
 1715 1720 1725  
 Ser Ser Ser Thr Asp Arg Glu Ser Glu Tyr Ser Phe Asn Gly Asp Ser  
 1730 1735 1740  
 Cys Ser Gly Leu Ser Ser Pro Ala Ser Pro Gly Thr Val Ser Pro Pro  
 1745 1750 1755 1760  
 Asn Glu Lys Val Ile Gln Ile His Glu Asn Gly Thr Met Lys Glu Ile  
 1765 1770 1775  
 Arg Ala Ile Ile Ala Asp Glu Ile Gly Val Ser Ala Asp Glu Ile Lys  
 1780 1785 1790  
 Ser Asp Glu Asn Leu Asn Glu Leu Gly Met Asp Ser Leu Leu Ser Leu  
 1795 1800 1805  
 Thr Val Leu Gly Lys Ile Arg Glu Ser Leu Asp Met Asp Leu Pro Gly  
 1810 1815 1820  
 Glu Phe Phe Ile Glu Asn Gln Thr Leu Asp Gln Ile Glu Thr Ala Leu  
 1825 1830 1835 1840  
 Asp Leu Lys Pro Lys Ala Val Pro Thr Ala Val Pro Gln Ser Gln Pro  
 1845 1850 1855  
 Ile Thr Leu Pro Gln Ser Gln Ser Thr Lys Gln Leu Ser Thr Arg Pro  
 1860 1865 1870  
 Thr Ser Ser Ser Asp Asn His Pro Pro Ala Thr Ser Ile Leu Leu Gln  
 1875 1880 1885  
 Gly Asn Pro Arg Thr Ala Ser Lys Thr Leu Phe Leu Phe Pro Asp Gly  
 1890 1895 1900

53

Ser Gly Ser Ala Thr Ser Tyr Ala Thr Ile Pro Gly Val Ser Pro Asn  
1905 1910 1915 1920  
Val Ala Val Tyr Gly Leu Asn Cys Pro Tyr Met Lys Ala Pro Glu Lys  
1925 1930 1935  
Leu Thr Cys Ser Leu Asp Ser Leu Thr Thr Pro Tyr Leu Ala Glu Ile  
1940 1945 1950  
Arg Arg Arg Gln Pro Thr Gly Pro Tyr Asn Leu Gly Gly Trp Ser Gln  
1955 1960 1965  
Ala Gly Ser Ala His Thr Thr Arg His Ala Ser Ser Tyr Cys Ser Arg  
1970 1975 1980

Ala Lys  
1985

<210> 14  
<211> 53  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 14  
atgaagcttg gggcttgagg gccaatggaa cgaaactagt gtaccacttg acc

53

<210> 15  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 15  
gacagatctg gcgccattcg ccattcag

28

<210> 16  
<211> 20  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer

<400> 16  
ggaatcggtc aatacactac

20

<210> 17  
<211> 33  
<212> DNA  
<213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 17

tgtagatctc tattcctttg ccctcggacg agt

33

&lt;210&gt; 18

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 18

ggccgcccacg gatattcttgg ccaaagaatt cctgg

35

&lt;210&gt; 19

&lt;211&gt; 35

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 19

cggtgcctat agaaccggtt tcttaaggac cgcgc

35

&lt;210&gt; 20

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 20

gayccmgtty ttyaayatg

19

&lt;210&gt; 21

&lt;211&gt; 17

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 21

gtccgtcert gcatytc

17

&lt;210&gt; 22

&lt;211&gt; 34

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 22

ataagaatgc ggccgcaatg gccctcgaaa cagc

34

&lt;210&gt; 23

&lt;211&gt; 29

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 23

aaatggcgcg ccgcgcccag aatgacacc

29

&lt;210&gt; 24

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 24

tgccacctgt agtctgcaat cag

23

&lt;210&gt; 25

&lt;211&gt; 24

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 25

tgactaaccg tgacaacttc gctg

24

&lt;210&gt; 26

&lt;211&gt; 19

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 26

ccaggatccg actgctcag

19

&lt;210&gt; 27

&lt;211&gt; 21

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

56

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 27

ctacatcgag atgcacggca c

21

&lt;210&gt; 28

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 28

ngtcgaswga nawgaa

16

&lt;210&gt; 29

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 29

gtncgaswca nawgtt

16

&lt;210&gt; 30

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 30

wgtgnagwan canaga

16

&lt;210&gt; 31

&lt;211&gt; 15

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 31

ntcgastwts gwggtt

15

&lt;210&gt; 32

&lt;211&gt; 16

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 32 16  
 tgwgnagwan casaga

<210> 33  
 <211> 16  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 33 16  
 agwgnagwan cawagg

<210> 34  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 34 14  
 cawcgngaa sgaa

<210> 35  
 <211> 14  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 35 14  
 tcstcgnaact wgga

<210> 36  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Description of Artificial Sequence: Primer

<400> 36 23  
 ttgttactgg agaggtaatg aag

<210> 37  
 <211> 23  
 <212> DNA  
 <213> Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 37

tgagacagat ctgcgagcc etc

23

&lt;210&gt; 38

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 38

atgtctccaa aggaagctga gc

22

&lt;210&gt; 39

&lt;211&gt; 22

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 39

tcgagtgatg gatactgctt cg

22

&lt;210&gt; 40

&lt;211&gt; 29

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 40

cggctacact agaaggacag tatttggtta

29

&lt;210&gt; 41

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 41

gtcaggcaac tatggatgaa cgaaatagac

30

&lt;210&gt; 42

&lt;211&gt; 25

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 42

acccatctca taataaacgt catgc

25

&lt;210&gt; 43

&lt;211&gt; 23

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 43

caactctatc agagcttggt tga

23

&lt;210&gt; 44

&lt;211&gt; 30

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 44

cccgaattca tgagctttgt tcaaataagg

30

&lt;210&gt; 45

&lt;211&gt; 39

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 45

ttattctaga ttttccatgg gaatggatac agtcttacg

39

&lt;210&gt; 46

&lt;211&gt; 33

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

&lt;220&gt;

&lt;223&gt; Description of Artificial Sequence: Primer

&lt;400&gt; 46

cgccaccatg gtgagcaagg gcgaggagct gtt

33

&lt;210&gt; 47

&lt;211&gt; 39

&lt;212&gt; DNA

&lt;213&gt; Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Primer  
  
<400> 47  
tatgatctag agtcgcggcc gctttacttg tacagctcg 39  
  
<210> 48  
<211> 25  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Primer  
  
<400> 48  
gcttctaatac cgtactagtg gatca 25  
  
<210> 49  
<211> 25  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Primer  
  
<400> 49  
ctttgatctt ttctacgggg tctga 25

**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☒ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**